

3

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
22 février 2001 (22.02.2001)

PCT

(10) Numéro de publication internationale
WO 01/12806 A2

(51) Classification internationale des brevets⁷: C12N 15/12,
C07K 14/47, C12Q 1/68, C12N 15/11, 5/10, C07K 16/18,
G01N 33/68, A61K 7/32, 7/46, A23C 9/152, C07K 19/00
// A61P 3/04, 3/06, 35/00

GACHON, Françoise [FR/FR]; 11, rue des Paillards,
F-63540 Romagnat (FR).

(21) Numéro de la demande internationale:

PCT/FR00/02319

(22) Date de dépôt international: 11 août 2000 (11.08.2000)

(25) Langue de dépôt: français

(26) Langue de publication: français

(30) Données relatives à la priorité:
99/10439 12 août 1999 (12.08.1999) FR

(71) Déposant (pour tous les États désignés sauf US): UNI-
VERSITE D'Auvergne [FR/FR]; 49, boulevard
François-Mitterrand, F-63000 Clermont-Ferrand (FR).

(71) Déposant et

(72) Inventeur: PITIOT, Gilles [FR/FR]; 151, rue du
Chevaleret, F-75013 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement): LA-
CAZETTE, Eric [FR/FR]; Appartement 211, 66,
boulevard La Fayette, F-63000 Clermont-Ferrand (FR).

(74) Mandataires: MARTIN, Jean-Jacques etc.; Cabinet
Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).

(81) États désignés (national): AE, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE,
DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,
NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) États désignés (régional): brevet ARIPO (GH, GM, KE,
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée:

— Sans rapport de recherche internationale, sera republiée
dès réception de ce rapport.

En ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abréviations" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: ODORANT-BINDING HUMAN PROTEINS FIXING HYDROPHOBIC LIGANDS: POLYPEPTIDES AND
POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND USES THEREOF

WO 01/12806 A2 (54) Titre: "ODORANT-BINDING" PROTEINES HUMAINES FIXANT DES LIGANDS HYDROPHOBES: POLYPEPTIDES
ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES ET LEURS APPLICATIONS

(57) Abstract: The invention concerns novel polypeptides fixing hydrophobic ligands and in particular odours, called OPBII (Odo-
rant Binding Protein), the polynucleotides coding for said polypeptides and specific antibodies directed against said polypeptides.
The invention also concerns the uses of said molecules for personal hygiene, uses in agri-food system and nutritional and therapeutic
uses

(57) Abrégé: L'invention concerne de nouveaux polypeptides fixant des ligands hydrophobes et notamment des odeurs, dénommés
OBPII (Odorant Binding Protein), les polynucléotides codant pour lesdits polypeptides et des anticorps spécifiques dirigés contre
lesdits polypeptides. L'invention concerne également les applications de ces molécules notamment pour l'hygiène corporelle, les
applications agro-alimentaires, nutritionnelles et thérapeutiques.



« ODORANT-BINDING » PROTEINES HUMAINES FIXANT DES
LIGANDS HYDROPHOBES : POLYPEPTIDES ET
POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES ET LEURS
APPLICATIONS ».

5 La présente invention concerne la mise en évidence de nouvelles protéines humaines fixant les odeurs, ci-après dénommées « OBP » (Odorant Binding Proteins) ainsi que leurs applications, tant au niveau thérapeutique que non thérapeutique.

10 La présente invention repose sur l'identification d'une famille de gènes de lipocalines composée de trois gènes et de deux pseudogènes sur le chromosome humain 9q34 ; les trois gènes correspondent au gène LCN1 déjà décrit et à deux nouveaux gènes faisant l'objet de la présente invention et dénommés hOBPIIa et hOBPIIb. De plus, l'invention repose sur l'attribution de nouvelles
15 fonctions à des lipocalines humaines déjà connues par la mise en évidence de nouveaux territoires d'expression.

 Bien qu'un certain nombre d'OBP ait déjà été mis en évidence (Pelosi *et al.*, 1996), telle l'OBP de rat par exemple (voir notamment le brevet EP-0 335 654), les OBP selon la présente
20 invention sont des OBP humaines qui présentent un très grand nombre d'avantages, comme cela ressortira de la suite du texte, par rapport aux protéines murines.

 Ces protéines OBP de la famille des lipocalines ont, pour certaines d'entre elles, été mentionnées indirectement dans le
25 brevet WO 99 07740, mais leur fonction d'OBP n'a jamais été décrite jusqu'à présent ; il en est ainsi également des protéines LCN1, rétinol-binding-protein (RBP) et Apolipoprotéine D (ApoD), comme cela sera explicité plus complètement ci-après.

Historiquement, la famille des lipocalines (Pervaiz et Brew, 1987) a été définie à partir de la protéine humaine fixant le rétinol (RBP) et à partir de 3 autres protéines : la β -lactoglobuline de boeuf, l' α 2 μ -globuline de rat et l' α 1-microglobuline humaine. A partir de ces protéines de référence et en utilisant des homologies de séquence, la famille des lipocalines s'est enrichie pour incorporer maintenant un grand nombre de protéines, plus d'une centaine, aussi bien chez les eucaryotes que chez les procaryotes (Flower *et al.*, 1995, 1996). Cette famille consiste en de petites protéines (160-190 acides aminés) contenant une poche hydrophobe et qui sont généralement sécrétées (Bocskai *et al.*, 1992 ; Senoo *et al.*, 1990, Zeng *et al.*, 1996 ; Miller, 1998), bien que dans certains cas il s'agisse de protéines qui restent associées à la membrane (Nagata *et al.*, 1991). Chez les vertébrés, les identités de séquence entre les différentes lipocalines tournent autour de 20%, toutefois, les identités de séquences sont plus importantes pour les protéines orthologues, de même que pour les récents gènes paralogues décrits (Igarashi *et al.*, 1992 ; Dewald *et al.*, 1992).

La présente invention concerne les polypeptides isolés OBPII codés par les deux nouveaux gènes humains OBPIIa et OBPIIb isolés localisés au locus 9q34 ; l'invention concerne également les séquences polynucléotidiques correspondantes, les ARNm correspondants, ainsi que les séquences régulatrices promotrices qui déterminent le profil d'expression dans les différents tissus et notamment dans les tissus sécréteurs. Ces deux gènes codent pour au moins sept polypeptides différents étant donné l'existence d'épissage alternatif des transcrits. Le gène OBPIIa code pour au moins 4 polypeptides différents dénommés OBPII $_{a\alpha}$, OBPII $_{a\beta}$, OBPII $_{a\gamma}$, OBPII $_{a\delta}$ et le gène OBPIIb code pour trois polypeptides

différents dénommés OBPII_{ba}, OBPII_{bβ}, OBPII_{bδ}. Même si un gène est principalement exprimé, toutes les formes de messagers sont retrouvées dans la structure nasale.

La présente invention a donc pour objet un polypeptide isolé
5 comprenant une séquence en amino-acides ayant au moins 90% d'identité avec les séquences en amino-acides SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14 et LCN1, ApoD et RBP.

Il doit être compris que l'invention concerne les polypeptides
10 obtenus par purification à partir de sources naturelles ou bien obtenus par les techniques recombinantes, comme cela sera décrit dans ce qui va suivre; l'invention concerne également les polypeptides obtenus par synthèse chimique qui peuvent alors comporter des acides aminés non naturels. Dans la présente
15 description, on utilisera le terme polypeptide pour désigner également une protéine ou un peptide.

Par pourcentage d'identité entre les séquences, on entend désigner le pourcentage d'acides aminés identiques entre les séquences obtenu avec le meilleur alignement de séquences
20 possibles. Ce pourcentage étant purement statistique et les différences entre les polypeptides étant réparties au hasard et sur toute la longueur.

La présente invention concerne également un polypeptide isolé caractérisé en ce qu'il comprend un polypeptide choisi parmi :

25 a) un polypeptide de séquence SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14;

b) un polypeptide variant de polypeptide de séquences d'acides aminés défini en a) ;

c) un polypeptide homologue au polypeptide défini en a) ou b) et comportant au moins 90 % d'identité avec ledit polypeptide de a) ;

5 d) un fragment d'au moins 15 acides aminés consécutifs d'un polypeptide défini en a), b) ou c) ;

e) un fragment biologiquement actif d'un polypeptide défini en a), b) ou c).

On entendra par polypeptide variant l'ensemble des polypeptides mutés pouvant exister naturellement, en particulier
10 chez l'être humain, et qui correspondent notamment à des troncatures, substitutions, délétions et/ou additions de résidus d'acides aminés. Les polypeptides variants selon l'invention conservent au moins un domaine de fixation à un ligand hydrophobe.

15 Par polypeptide homologue, on entendra désigner les polypeptides présentant, par rapport aux polypeptides de séquence SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 et SEQ ID N°14, certaines modifications comme en particulier une délétion, addition ou substitution d'au moins un
20 acide aminé, une troncature, un allongement et/ou une fusion chimérique. Parmi les polypeptides homologues, on préfère ceux dont la séquence d'acides aminés présente au moins 90 % d'identité, de préférence 95 %, de manière préférée 97 %, et de manière encore préférée 99 % d'identité avec les séquences d'acides
25 aminés des polypeptides selon l'invention. Dans le cas d'une substitution, un ou plusieurs acides aminés consécutifs ou non consécutifs, sont remplacés par des « acides aminés équivalents ». L'expression « acide aminé équivalent » vise ici à désigner tout acide aminé susceptible d'être substitué à l'un des acides aminés de la

structure de base sans cependant modifier les caractéristiques ou propriétés fonctionnelles essentielles des polypeptides correspondants, leurs activités biologiques, telles par exemple l'induction *in vivo* d'anticorps capables de reconnaître le polypeptide dont la séquence d'acides aminés est comprise dans la séquence d'acides aminés SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14, ou l'un de ses fragments. Ces acides aminés équivalents peuvent être déterminés soit en s'appuyant sur leur homologie de structure avec les acides aminés auxquels ils se substituent, soit sur les résultats des essais d'activité biologique croisée auxquels les différents polypeptides sont susceptibles de donner lieu. A titre d'exemple, on mentionnera les possibilités de substitutions susceptibles d'être effectuées sans qu'il en résulte une modification approfondie des activités biologiques des polypeptides modifiés correspondants, les remplacements, par exemple, de la leucine par la valine ou l'isoleucine, de l'acide aspartique par l'acide glutamique, de la glutamine par l'asparagine, de l'arginine par la lysine etc., les substitutions inverses étant naturellement envisageables dans les mêmes conditions. Ainsi, on peut envisager d'introduire certaines modifications comme en particulier une délétion, addition ou substitution d'au moins un acide aminé au niveau des hélices alpha de la protéine sans détruire le calice formé par la structure composée des feuillets beta ; de la même manière, il est possible d'introduire des amino-acides équivalents qui permettent de conserver aux feuillets bêta le caractère hydrophobe. Il peut également être intéressant d'introduire des modifications dans la séquence des polypeptides de l'invention pour générer des polypeptides homologues dépourvus de sites protéasiques.

Par fragment biologiquement actif, on entendra désigner en particulier un fragment de séquence d'acides aminés de polypeptide selon l'invention présentant au moins une des caractéristiques ou propriétés fonctionnelles des polypeptides selon l'invention, notamment en ce que : (i) il est capable d'être reconnu par un anticorps spécifique d'un polypeptide selon l'invention ou bien par des anticorps produits par des patients au cours d'une réaction immunitaire; (ii) il présente au moins l'un des domaines ou régions tels que définis ci-après ; (iii) il est capable de lier un ligand hydrophobe et notamment des molécules odorantes, de préférence les phéromones ; (iv) il est capable de lier spécifiquement un récepteur.

Par fragment de polypeptide, on entend désigner un polypeptide comportant au minimum 15 acides aminés, de préférence 18 acides aminés, de manière préférée 25 et de manière encore préférée 50 acides aminés. Les fragments de polypeptide selon l'invention obtenus par clivage dudit polypeptide par une enzyme protéolytique, par un réactif chimique, ou encore en plaçant ledit polypeptide dans un environnement très acide font également partie de l'invention. Lorsque l'on souhaite utiliser une séquence de 15, 18, 25 ou de 50 amino-acides, il s'agit bien entendu, de préférence, des parties correspondant à des épitopes fonctionnels des polypeptides précédents, lesquels pourront être intégrés dans un polypeptide ayant une structure plus longue sous forme par exemple de protéine de fusion, ceci dépendra des applications que l'on souhaite de ces polypeptides.

Selon un mode préféré de réalisation, la présente invention concerne un polypeptide isolé sélectionné parmi un polypeptide correspondant à la séquence SEQ ID N°2 et dénommé OBPII_{aa} , à la

séquence SEQ ID N°4 et dénommé OBPII_{aβ}, à la séquence SEQ ID N°6 et dénommé OBPII_{aγ}, à la séquence SEQ ID N°10 et dénommé OBPII_{ba}, à la séquence SEQ ID N°12 et dénommé OBPII_{bβ}.

Les polypeptides selon l'invention se caractérisent en ce qu'ils comportent de préférence le domaine Gly-Thr-Trp-Tyr.

L'invention concerne également le polynucléotide isolé caractérisé en ce qu'il code pour un polypeptide tel que décrit précédemment. Le polynucléotide selon l'invention est choisi parmi le polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13.

L'invention porte également sur un polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

a) un polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13 ou dont la séquence est celle de l'ARN correspondant à la séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13 ;

b) un polynucléotide dont la séquence est complémentaire de la séquence d'un polynucléotide défini en a),

c) un polynucléotide dont la séquence comporte au moins 80% d'identité, de préférence 90 %, de manière préférée 95 %, et de manière encore préférée 97 % d'identité avec un polynucléotide défini en a) ou b),

d) un polynucléotide hybridant dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b) ou c),

e) un fragment d'au moins 15 nucléotides consécutifs, de préférence 21 nucléotides consécutifs, et de manière préférée 30 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).

Dans la présente description, on entendra désigner par polynucléotide, oligonucléotide, séquence de polynucléotide, séquence nucléotidique ou acide nucléique un fragment d'ADN, aussi bien un ADN double brin, un ADN simple brin que des
5 produits de transcription desdits ADN, et/ou un fragment d'ARN, lesdits fragments naturels isolés, ou de synthèse, comportant ou non des nucléotides non naturels, désignant un enchaînement précis de nucléotides, modifiés ou non, permettant de définir un fragment ou une région d'un acide nucléique.

10 Par polynucléotide de séquence complémentaire, on entend désigner tout ADN dont les nucléotides sont complémentaires de ceux de la SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11, SEQ ID N° 13 ou d'une partie de la SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ
15 ID N°11, SEQ ID N° 13 et dont l'orientation est inversée.

Par pourcentage d'identité au sens de la présente invention, on entend un pourcentage de bases identiques entre les polynucléotides obtenus après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les
20 deux polynucléotides étant réparties au hasard et sur toute leur longueur.

Une hybridation dans des conditions de forte stringence signifie au sens de la présente invention, que les conditions de température et de force ionique sont choisies de telle manière
25 qu'elles permettent le maintien de l'hybridation entre deux fragments d'ADN complémentaires. A titre illustratif, des conditions de forte stringence de l'étape d'hybridation aux fins de définir les fragments polynucléotidiques décrits ci-dessus, sont avantageusement les suivantes :

L'hybridation ADN-ADN ou ADN-ARN est réalisée en deux étapes: (1) préhybridation à 42°C pendant 3 heures en tampon phosphate (20 mM, pH 7,5) contenant 5 x SSC (1 x SSC correspond à une solution 0,15 M NaCl + 0,015 M citrate de sodium), 50 % de formamide, 7 % de sodium dodécyl sulfate (SDS), 10 x Denhard's, 5 % de dextran sulfate et 100 µg/ml d'ADN de sperme de saumon ; (2) hybridation proprement dite pendant 20 heures à une température dépendant de la taille de la sonde (i.e. : 42°C, pour une sonde de taille > 100 nucléotides) suivie de 2 lavages de 20 minutes à 20°C en 2 x SSC + 2 % SDS, 1 lavage de 20 minutes à 20°C en 0,1 x SSC + 0,1 % SDS. Le dernier lavage est pratiqué en 0,1 x SSC + 0,1 % SDS pendant 30 minutes à 60°C pour une sonde de taille > 100 nucléotides. Les conditions d'hybridation de forte stringence décrites ci-avant pour un polynucléotide de taille définie, seront adaptées par l'homme du métier pour des oligonucléotides de taille plus grande ou plus petite, selon l'enseignement de Sambrook *et al.*(1989).

Avantageusement, un fragment nucléotidique répondant à la définition précédente aura au moins 15 nucléotides consécutifs, de préférence au moins 21 nucléotides, et encore plus préférentiellement au moins 30 nucléotides consécutifs de la séquence dont il est issu.

Selon un mode de réalisation de l'invention, le polynucléotide selon l'invention se caractérise en ce qu'il est marqué directement ou indirectement par un composé radioactif ou un composé non radioactif. Le polynucléotide selon l'invention est utilisé en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques ; l'invention porte également sur l'utilisation d'un polynucléotide selon l'invention en tant que sonde

pour la détection de séquences nucléiques. Selon l'invention, les fragments de polynucléotides pouvant être utilisés comme sonde ou comme amorce dans des procédés de détection, d'identification, de dosage ou d'amplification de séquence nucléique, présenteront une
5 taille minimale de 9 bases, de préférence de 18 bases, et de manière plus préférée 36 bases. Enfin, l'invention porte sur l'utilisation d'un polynucléotide selon l'invention en tant que oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant en l'occurrence un polypeptide selon l'invention.

10 Les séquences de polynucléotides selon l'invention non marquées peuvent être utilisées directement comme sonde, amorce ou oligonucléotide ; cependant les séquences utilisées sont généralement marquées pour obtenir des séquences utilisables pour de nombreuses applications. Le marquage des amorces, des
15 sondes, des oligonucléotides selon l'invention est réalisé par des éléments radioactifs ou par des molécules non radioactives ; parmi les isotopes radioactifs utilisés, on peut citer le ^{32}P , le ^{33}P , le ^{35}S , le ^3H ou le ^{125}I . Les entités non radioactives sont sélectionnées parmi les ligands tels la biotine, l'avidine, la streptavidine, la dioxygénine,
20 les haptènes, les colorants, les agents luminescents tels que les agents radioluminescents, chimiluminescents, bioluminescents, fluorescents, phosphorescents.

L'invention comprend également une méthode de détection et/ou de dosage d'acide nucléique selon l'invention, dans un
25 échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i) d'isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ; (ii) d'amplification spécifique de l'ADN

codant pour le polypeptide selon l'invention à l'aide d'amorces ; (iii) d'analyse des produits d'amplification.

L'invention comprend en outre un nécessaire pour la détection et/ou le dosage d'un acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) un couple d'amorces nucléiques selon l'invention, (ii) les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN, et éventuellement (iii) un composant permettant de vérifier la séquence du fragment amplifié, plus particulièrement une sonde selon l'invention.

L'invention comprend aussi une méthode de détection et/ou de dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i) de mise en contact d'une sonde selon l'invention avec un échantillon biologique ; (ii) de détection et/ou de dosage de l'hybride formé entre ladite sonde et l'ADN de l'échantillon biologique.

L'invention comprend également un nécessaire pour la détection et/ou le dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) une sonde selon l'invention, (ii) les réactifs nécessaires à la mise en oeuvre d'une réaction d'hybridation, et le cas échéant, (iii) un couple d'amorces selon l'invention, ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN.

L'invention concerne particulièrement les procédés selon l'invention et décrits ci-dessus, pour la détection et le diagnostic de cellules d'origine cancéreuse et principalement cancers du sein, de l'utérus, de l'ovaire, de la prostate et du poumon.

Les polynucléotides selon l'invention peuvent ainsi être utilisés comme amorce et/ou sonde dans des procédés mettant en

oeuvre notamment la technique PCR (réaction en chaîne à la polymérase)(Erlich, 1989 ; Innis *et al.*, 1990, et Rolfs *et al.*, 1991). Cette technique nécessite le choix de paires d'amorces oligonucléotidiques encadrant le fragment qui doit être amplifié. On
5 peut, par exemple, se référer à la technique décrite dans le brevet américain U.S. N° 4 683 202. Les fragments amplifiés peuvent être identifiés, par exemple après une électrophorèse en gel d'agarose ou de polyacrylamide, ou après une technique chromatographique
10 comme la filtration sur gel ou la chromatographie échangeuse d'ions. La spécificité de l'amplification peut être contrôlée par hybridation moléculaire en utilisant comme sonde les séquences nucléotidiques de polynucléotides de l'invention, de plasmides contenant ces séquences ou leurs produits d'amplification. Les
15 fragments nucléotidiques amplifiés peuvent être utilisés comme réactifs dans des réactions d'hybridation afin de mettre en évidence la présence, dans un échantillon biologique, d'un acide nucléique cible de séquence complémentaire à celle desdits fragments nucléotidiques amplifiés.

L'invention vise également les fragments nucléotidiques
20 susceptibles d'être obtenus par amplification à l'aide d'amorces selon l'invention.

D'autres techniques d'amplification de l'acide nucléique cible peuvent être avantageusement employées comme alternative à la PCR (PCR-like) à l'aide de couple d'amorces de séquences
25 nucléotidiques selon l'invention. Par PCR-like on entendra désigner toutes les méthodes mettant en oeuvre des reproductions directes ou indirectes des séquences d'acides nucléiques, ou bien dans lesquelles les systèmes de marquage ont été amplifiés, ces techniques sont bien entendu connues, en général il s'agit de

l'amplification de l'ADN par une polymérase ; lorsque l'échantillon d'origine est un ARN il convient préalablement d'effectuer une transcription inverse. Il existe actuellement de très nombreux procédés permettant cette amplification, comme par exemple la

5 technique SDA (Strand Displacement Amplification) ou technique d'amplification à déplacement de brin (Walker *et al.*, 1992), la technique TAS (Transcription-based Amplification System) décrite par Kwoh *et al.* en 1989, la technique 3SR (Self-Sustained Sequence Replication) décrite par Guatelli *et al.* en 1990, la

10 technique NASBA (Nucleic Acid Sequence Based Amplification) décrite par Kievitis *et al.* en 1991, la technique TMA (Transcription Mediated Amplification), la technique LCR (Ligase Chain Reaction) décrite par Landegren *et al.* en 1988 et perfectionnée par Barany *et al.* en 1991, qui emploie une ligase thermostable, la technique de

15 RCR (Repair Chain Reaction) décrite par Segev en 1992, la technique CPR (Cycling Probe Reaction) décrite par Duck *et al.* en 1990, la technique d'amplification à la Q-béta-réplase décrite par Miele *et al.* en 1983 et perfectionnée notamment par Chu *et al.* en 1986 et Lizardi *et al.* en 1988, puis par Burg *et al.* (1996) ainsi que

20 par Stone *et al.* (1996).

Dans le cas où le polynucléotide cible à détecter est un ARN, par exemple un ARNm, on utilisera avantageusement, préalablement à la mise en oeuvre d'une réaction d'amplification à l'aide des amorces selon l'invention ou à la mise en oeuvre d'un

25 procédé de détection à l'aide des sondes de l'invention, une enzyme de type transcriptase inverse afin d'obtenir un ADNc à partir de l'ARN contenu dans l'échantillon biologique. L'ADNc obtenu servira alors de cible pour les amorces ou les sondes mises en oeuvre dans le procédé d'amplification ou de détection selon l'invention.

Les sondes nucléotidiques selon l'invention hybrident spécifiquement avec une molécule d'ADN ou d'ARN de polynucléotide selon l'invention, plus particulièrement avec les séquences SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, 5 SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13, dans des conditions d'hybridation de forte stringence telles que données sous forme d'exemple précédemment.

La technique d'hybridation peut être réalisée de manières diverses (Matthews *et al.*, 1988). La méthode la plus générale 10 consiste à immobiliser l'acide nucléique extrait des cellules de différents tissus ou de cellules en culture sur un support (tel que la nitrocellulose, le nylon, le polystyrène) et à incuber, dans des conditions bien définies, l'acide nucléique cible immobilisé avec la sonde. Après l'hybridation, l'excès de sonde est éliminé et les 15 molécules hybrides formées sont détectées par la méthode appropriée (mesure de la radioactivité, de la fluorescence ou de l'activité enzymatique liée à la sonde).

Selon un autre mode de mise en oeuvre des sondes nucléiques selon l'invention, ces dernières peuvent être utilisées 20 comme sonde de capture. Dans ce cas, une sonde, dite « sonde de capture », est immobilisée sur un support et sert à capturer par hybridation spécifique l'acide nucléique cible obtenu à partir de l'échantillon biologique à tester et l'acide nucléique cible est ensuite détecté grâce à une seconde sonde, dite « sonde de détection », 25 marquée par un élément facilement détectable.

Dans un mode préféré de réalisation, l'invention comprend l'utilisation d'un oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant. Parmi les fragments d'acides nucléiques intéressants, il faut citer en

particulier les oligonucléotides anti-sens, c'est-à-dire dont la structure assure, par hybridation avec la séquence cible, une inhibition de l'expression du produit correspondant. Il faut également citer les oligonucléotides sens qui, par interaction avec
5 des protéines impliquées dans la régulation de l'expression du produit correspondant, induiront soit une inhibition, soit une activation de cette expression. Les oligonucléotides selon l'invention présentent une taille minimale de 9 bases, de préférence de 18 bases, et de manière plus préférée 36 bases.

10 L'invention concerne un vecteur recombinant de clonage d'un polynucléotide selon l'invention et/ou d'expression d'un polypeptide selon l'invention caractérisé en ce qu'il contient un polynucléotide selon l'invention tel que précédemment décrit. Le vecteur selon l'invention est caractérisé en ce qu'il comporte les
15 éléments permettant l'expression desdites séquences dans une cellule hôte et éventuellement la sécrétion desdites séquences hors de la cellule hôte. Par « vecteur d'expression », on entend aussi bien des vecteurs d'expression à réplication autonome du type plasmide que des systèmes destinés à assurer l'intégration dans les cellules,
20 mais ces vecteurs d'expression pourront être également des vecteurs d'expression de type viral ou bien même, lorsque l'on souhaite réaliser par exemple de la thérapie génique, de l'ADN nu. Parmi les vecteurs viraux, on préfère ceux dérivés de l'adénovirus, du virus associé à l'adénovirus (AAV), de rétrovirus, des lentivirus,
25 et de préférence les dérivés de HIV, des poxvirus, du virus de l'herpes pour l'expression en système eucaryote. Parmi les vecteurs non viraux, on préfère les polynucléotides nus tels que l'ADN nu ou l'ARN nu selon la technique développée par la société VICAL, les chromosomes artificiels de levure (YAC, yeast artificial

chromosome) pour l'expression dans la levure, les chromosomes artificiels de souris (MAC, mouse artificial chromosome) pour l'expression dans les cellules murines et de manière préférée les chromosomes artificiels d'homme (HAC, human artificial chromosome) pour l'expression dans les cellules humaines.

Selon un mode particulier de réalisation, le vecteur selon l'invention comporte des éléments de contrôle de l'expression des polypeptides ; ces éléments de contrôle sont choisis de préférence parmi (i) la séquence promotrice du gène hOBPIIa selon l'invention qui correspond à la séquence SEQ ID N°15 et/ou parmi la séquence promotrice du gène hOBPIIb selon l'invention qui correspond à la séquence SEQ ID N°16 ; (ii) un polynucléotide dont la séquence est complémentaire à la séquence SEQ ID N° 15 et SEQ ID N° 16 ; (iii) un polynucléotide dont la séquence comporte au moins 80% d'identité avec un polynucléotide défini en (i) ou (ii) ; (iv) un polynucléotide hybridant dans des conditions de forte stringence avec la séquence de polynucléotide définie en (i), (ii), (iii). Les outils informatiques à la disposition de l'homme du métier lui permettent aisément d'identifier les boîtes régulatrices promotrices nécessaires et suffisantes au contrôle de l'expression génique, notamment les boîtes TATA, CCAAT, GC, ainsi que les séquences régulatrices stimulatrices (« enhancer ») ou inhibitrices (« silencers ») qui contrôlent en CIS l'expression des gènes selon l'invention ; parmi ces séquences régulatrices, il convient de citer l'I_{RE}, MRE, CRE.

Il est également dans l'étendue de l'invention d'utiliser les éléments ci-dessus définis et choisis parmi la séquence SEQ ID N°15 et/ou SEQ ID N°16 pour contrôler l'expression de polypeptides hétérologues autres que ceux de l'invention et

notamment pour diriger l'expression de polypeptides hétérologues dans les types cellulaires dans lesquels les polypeptides selon l'invention s'expriment normalement.

L'invention comprend en outre les cellules hôtes, 5 notamment les cellules eucaryotes et procaryotes, caractérisées en ce qu'elles sont transformées par les vecteurs selon l'invention. De préférence, les cellules hôtes sont transformées dans des conditions permettant l'expression d'un polypeptide recombinant selon l'invention. L'hôte cellulaire peut être choisi parmi les cellules 10 bactériennes mais également parmi les cellules de levure, de même que parmi les cellules végétales et animales ; de préférence, l'hôte cellulaire est une cellule de mammifères (Edwards et Aruffo, 1993), mais également une cellule d'insectes dans laquelle on peut utiliser des procédés mettant en oeuvre des baculovirus par exemple 15 (Luckow, 1993). Ces cellules peuvent être obtenues par l'introduction dans des cellules hôtes d'une séquence nucléotidique insérée dans un vecteur tel que défini ci-dessus, puis la mise en culture desdites cellules dans des conditions permettant la réplication et/ou l'expression de la séquence nucléotidique 20 transfectée.

Selon un mode particulier de réalisation, l'invention porte également sur un animal ou une plante transgénique qui comporte des cellules hôtes selon l'invention.

L'invention concerne également une méthode de préparation 25 d'un polypeptide caractérisé en ce qu'il met en oeuvre un vecteur selon l'invention. Plus particulièrement, l'invention porte sur une méthode de préparation d'un polypeptide recombinant caractérisé en ce que l'on cultive des cellules transformées selon l'invention

dans des conditions permettant l'expression dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.

Le polypeptide selon l'invention est susceptible d'être obtenu selon un procédé de l'invention et selon les techniques de production de polypeptides recombinants connues de l'homme du métier. La présente invention concerne donc le polypeptide recombinant susceptible d'être obtenu par la méthode ci-dessus présentée. Dans ce cas, la séquence d'acide nucléique utilisée est placée sous le contrôle de signaux permettant son expression dans un hôte cellulaire. Un système efficace de production d'un polypeptide recombinant nécessite de disposer d'un vecteur, par exemple d'origine plasmidique ou virale, et d'une cellule hôte compatible. Le vecteur doit comporter un promoteur, des signaux d'initiation et de terminaison de la traduction, ainsi que des régions appropriées de régulation de la transcription. Il doit pouvoir être maintenu de façon stable dans la cellule et peut éventuellement posséder des signaux particuliers spécifiant la sécrétion du polypeptide traduit. Ces différents signaux de contrôle sont choisis en fonction de l'hôte cellulaire utilisé. A cet effet, les séquences d'acide nucléique selon l'invention peuvent être insérées dans des vecteurs à réplication autonome au sein de l'hôte choisi, ou des vecteurs intégratifs de l'hôte choisi. De tels vecteurs seront préparés selon les méthodes couramment utilisées par l'homme du métier, et les clones en résultant peuvent être introduits dans un hôte approprié par des méthodes standards telles par exemple la transfection par précipitation au phosphate de calcium, la lipofection, l'électroporation, le choc thermique.

Les polypeptides recombinants obtenus comme indiqué ci-dessus, peuvent aussi bien se présenter sous forme glycosylée que

non glycosylée et peuvent présenter ou non la structure tertiaire naturelle.

Les polypeptides obtenus par synthèse chimique et pouvant comporter des acides aminés non naturels correspondant auxdits
5 polypeptides recombinants, sont également compris dans l'invention. Les peptides selon l'invention peuvent également être préparés par les techniques classiques, dans le domaine de la synthèse des peptides. Cette synthèse peut être réalisée en solution homogène ou en phase solide.

10 Les procédés de purification de polypeptide recombinant utilisés sont connus de l'homme du métier. Le polypeptide recombinant peut être purifié à partir de lysats et extraits cellulaires, du surnageant du milieu de culture, par des méthodes utilisées individuellement ou en combinaison, telles que le
15 fractionnement, les méthodes de chromatographie, les techniques d'immuno-affinité à l'aide d'anticorps mono- ou polyclonaux spécifiques, etc. Une variante préférée consiste à produire un polypeptide recombinant fusionné à une protéine « porteuse » (protéine chimère). L'avantage de ce système est qu'il permet une
20 stabilisation et une diminution de la protéolyse du produit recombinant, une augmentation de la solubilité au cours de la renaturation *in vitro* et/ou une simplification de la purification lorsque le partenaire de fusion possède une affinité pour un ligand spécifique.

25 La présente invention concerne l'utilisation d'un polypeptide choisi parmi un polypeptide hOBPIIa et hOBPIIb selon l'invention ou l'un de leur fragment et parmi LCN1, la « rétinol binding protein » (RBP) et l'apolipoprotéine D (ApoD) comme protéine de liaison à un ligand hydrophobe, de préférence une molécule

odorante. En effet, outre les polypeptides OBPII mentionnés précédemment, la fonction OBP des lipocalines LCN1, RBP et ApoD n'a jamais été suggérée. En effet, si la présence de LCN1 dans le mucus nasal a été décrite (Redl et al., 1992), c'est en fait une autre
5 fonction qui lui a été attribuée, à savoir une fonction d'anti-protéase, alors que la fonction principale de la protéine LCN1 semble bien être le transport de ligands hydrophobes, d'ailleurs les essais réalisés dans le cadre de la présente invention avec une protéine LCN1 recombinante n'ont pas permis de mettre en
10 évidence une activité anti-protéase. Il en va de même pour la « rétinol-binding protein » et l'apolipoprotéine D qui n'ont jamais été présentées comme ayant une fonction OBP.

La présente invention concerne également l'utilisation d'un polypeptide selon l'invention ou l'un de ses fragments, LCN1, RBP
15 et l'apolipoprotéine D comme inhibiteur compétitif, comme agoniste ou antagoniste des récepteurs cellulaires aux lipocalines. L'utilisation d'inhibiteur de liaison des lipocalines à leur récepteur peut être utilisée dans une stratégie anti-cancéreuse ; en effet, un certain nombre de tumeurs sont hormono-dépendantes : ainsi les
20 tumeurs du sein sont sensibles aux stéroïdes ; ces stéroïdes étant transportés par les lipocalines, il est dans l'étendue de l'invention de fournir des inhibiteurs de liaison aux polypeptides selon l'invention et LCN1, OBP ou APO-D pour éviter la fixation des hormones stéroïdes au niveau des récepteurs des cellules
25 tumorales.

L'invention concerne également un anticorps monoclonal ou polyclonal et ses fragments, caractérisés en ce qu'ils lient spécifiquement un polypeptide selon l'invention. Les anticorps chimériques, les anticorps humanisés et les anticorps simple

chaîne font également partie de l'invention. Les fragments d'anticorps selon l'invention sont de préférence des fragments Fab ou F(ab')₂.

Les polypeptides selon l'invention permettent de préparer
5 des anticorps monoclonaux ou polyclonaux. Les anticorps monoclonaux pourront avantageusement être préparés à partir d'hybridomes selon la technique décrite par Kohler et Milstein en 1975. Les anticorps polyclonaux pourront être préparés, par exemple par immunisation d'un animal, en particulier une souris,
10 avec un polypeptide selon l'invention associé à un adjuvant de la réponse immunitaire, puis purification des anticorps spécifiques contenus dans le sérum des animaux immunisés sur une colonne d'affinité sur laquelle a préalablement été fixé le polypeptide ayant servi d'antigène. Les anticorps polyclonaux selon l'invention
15 peuvent aussi être préparés par purification sur une colonne d'affinité, sur laquelle a préalablement été immobilisé un polypeptide selon l'invention.

L'invention porte également sur un anticorps monoclonal spécifique d'un polypeptide selon l'invention et capable d'inhiber
20 l'interaction entre ledit polypeptide et le récepteur cellulaire sur lequel se lie spécifiquement ledit polypeptide. Selon un autre mode de réalisation, l'anticorps monoclonal selon l'invention est capable d'inhiber l'interaction entre ledit polypeptide et ses ligands hydrophobes, de préférence les molécules odorantes et de manière
25 préférée les phéromones avec lesquelles ledit polypeptide se lie.

Les anticorps de l'invention pourront également être marqués de la même manière que décrit précédemment pour les sondes nucléiques de l'invention et de manière préférée avec un marquage de type enzymatique, fluorescent ou radioactif. De tels

anticorps marqués pourront être utilisés pour la détection de ces polypeptides dans un échantillon biologique. De préférence, l'échantillon biologique est constitué par un fluide par exemple du sérum, du sang ou des biopsies humaines. Ils constituent ainsi un
5 moyen d'analyse de l'expression de polypeptide selon l'invention, par exemple par immunofluorescence, marquage à l'or, immunoconjugués enzymatiques.

Plus généralement, les anticorps de l'invention peuvent être avantageusement mis en oeuvre dans toute situation où
10 l'expression d'un polypeptide selon l'invention doit être observée, et plus particulièrement en immunocytochimie, en immunohistochimie ou dans des expériences de « western blotting », dans les techniques ELISA et RIA. Il est ainsi dans l'étendue de l'invention de fournir une méthode de détection et/ou
15 de dosage d'un polypeptide selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes de mise en contact de l'échantillon biologique avec des anticorps selon l'invention puis de mise en évidence du complexe antigène-anticorps formé.

20 Entre également dans le cadre de l'invention, un nécessaire pour la détection et/ou le dosage d'un polypeptide selon l'invention dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) un anticorps monoclonal ou polyclonal tel que décrit précédemment ; (ii) le cas échéant, les réactifs pour la
25 constitution du milieu propice à la réaction immunologique ; (iii) les réactifs permettant la détection des complexes antigène-anticorps produits par la réaction immunologique. Ce nécessaire est notamment utile à la réalisation d'expériences de Western Blotting et aux expériences d'immunoprécipitation.

Les hOBPII selon la présente invention peuvent être utilisées dans de nombreuses applications.

Les premières applications des OBP résident dans le contrôle des odeurs ; ces applications concernent essentiellement
5 l'hygiène corporelle (parfumerie, cosmétologie, pharmacie) ou collective.

Elles peuvent être utilisées dans un procédé pour contrôler la volatilisation d'une odeur ; un tel procédé se caractérise en ce qu'il comprend une étape de liaison de ladite odeur avec un
10 polypeptide selon l'invention ou avec les protéines LCN1, RBP ou l'apolipoprotéine D.

Il est également possible de fixer l'odeur sur un support solide en utilisant une OBPII selon l'invention attachée audit support, cette fixation pouvant être aussi bien une fixation
15 covalente que non covalente, par exemple par adsorption ou par des fixations de type avidine-biotine, si cela est nécessaire. Dans ces conditions, on obtient un support parfumé retenant les odeurs plus longtemps, libérant progressivement les odeurs et pouvant être utilisé aussi bien dans le domaine cosmétique que dans le domaine
20 des produits d'entretiens en général. Les supports utilisés pourront être des supports de type plaque ou de type bille par exemple. Bien entendu, les OBPII selon la présente invention seront plus particulièrement utilisables dans l'industrie de la parfumerie et de la cosmétique où les OBP seront utilisées sous forme d'un mélange
25 liquide destiné à contrôler la volatilisation des odeurs après que la composition ait été répandue sur la peau humaine. Ceci permet de prolonger la « tenue » des parfums ou d'entrer dans la composition des déodorants corporels notamment. Bien entendu, le terme « parfum » englobe de façon générale les mélanges connus en

parfumerie, à base d'alcool ou sous forme aqueuse, et contenant notamment des huiles essentielles. Le polypeptide selon l'invention peut également entrer dans la composition de crèmes à base de rétinol en tant qu'agent de transport et de protection dudit rétinol
5 pour des applications cosmétologiques et notamment pour prévenir, effacer et traiter les rides, les ridules de la peau, lutter contre le relâchement cutané et/ou sous-cutané.

Parmi les applications concernant l'hygiène collective, le procédé selon l'invention peut être utilisé dans un dispositif visant
10 à désodoriser les locaux tels par exemple les animaleries, les étables. Un tel dispositif peut également être envisagé pour désodoriser le flux d'air entrant dans un appareil à air conditionné ; un tel dispositif serait très utile dans les zones géographiques polluées.

15 L'invention concerne également un procédé de criblage de molécule, de préférence les odeurs ou saveurs qui comprend de passer la molécule sur un substrat qui comprend un polypeptide selon l'invention ou, LCN1, RBP et l'apolipoprotéine D lié audit substrat, ledit polypeptide liant ladite odeur ou saveur et de
20 récupérer ladite odeur ou saveur depuis le polypeptide si nécessaire. Les OBP selon l'invention pourraient ainsi permettre l'isolement d'odeurs ou de parfums en les fixant sur un support tel que décrit précédemment mais qui, dans ce cas, pourra être par exemple une colonne de chromatographie, en faisant passer sur
25 celle-ci le produit sur lequel on souhaite capter les odeurs. Il est bien entendu possible, toujours dans ce même type d'application, d'effectuer une analyse d'un parfum complexe en utilisant les différences de rétention des différents produits par les OBP.

Le procédé d'isolement décrit précédemment peut également être utilisé pour séparer des « odeurs » particulières, notamment les phéromones humaines.

Les hOBPII selon la présente invention peuvent être utilisées
5 pour des applications agro-alimentaires ou industrielles.

Les OBP selon la présente invention de part leurs caractéristiques peuvent permettre de solubiliser certaines molécules lipophiles en les associant avec lesdits OBP. Ainsi les polypeptides selon l'invention peuvent être utilisés en combinaison
10 avec des acides gras alimentaires à titre d'additif alimentaire. La présente invention concerne un procédé pour solubiliser des molécules lipophiles caractérisé en ce qu'il comprend la liaison de ladite molécule lipophile à un polypeptide selon l'invention, à LCN1, RBP ou à l'apoD.

15 Les OBPII, de même que LCN1, RBP et ApoD sont susceptibles d'être impliquées dans le transport des acides gras et dans les mécanismes biologiques permettant la détection de la charge en acides gras de la ration alimentaire, au niveau buccal notamment. L'invention concerne donc également l'utilisation des
20 polypeptides précédents en association avec des acides gras pour diminuer la consommation d'acides gras, notamment dans les hyperlipidémies ou les obésités. Les polypeptides précédents peuvent donc être utilisés pour le traitement des hyperlipidémies et de l'obésité. En effet, ces protéines participent à la détection de la
25 teneur en acides gras dans la prise alimentaire (Gilbertson, 1998), un excès de ces protéines doit conduire à leurrer le système physiologique détectant la charge en gras d'une ration alimentaire. Ainsi, une portion alimentaire pauvre en graisse mais supplémentée en OBP ou ApoD ou RBP ou LCN1 préalablement

chargées en acide gras, sera faussement identifiée comme riche en graisse.

Une autre application est de compléter un lait non maternel avec un ou plusieurs des polypeptides décrits
5 précédemment.

Les OBP selon la présente invention peuvent être utilisées en thérapie préventive et curative.

Ainsi, l'absence de détection de ce type de protéines dans un prélèvement biologique en utilisant un anticorps, une amorce, une
10 sonde selon l'invention pourrait être un élément de diagnostic des anosmies.

De même, il est possible de prévoir la détection des anticorps anti-OBPII dans un échantillon biologique d'origine humaine, la présence ou le dosage de ces anticorps pourrait être en rapport
15 avec certains types d'allergies, notamment chez les asthmatiques. Dans ce cas, il pourrait être possible de traiter ce type d'allergie par administration de fragments de polypeptides tels que décrits, ce qui diminuerait les réactions immunitaires aux allergies externes. L'invention concerne donc un procédé de détection d'anticorps
20 dirigé contre les OBPII humaines (hOBPII) dans le sérum humain de patient allergique et/ou asthmatique en utilisant un polypeptide hOBPII selon l'invention.

Selon un autre mode de réalisation, l'invention concerne un procédé de détection d'anticorps dirigé contre les OBPII humaines
25 (hOBPII) dans un fluide biologique de patient atteint d'un cancer, et notamment du cancer de la prostate et/ou du cancer du sein et/ou du cancer de l'utérus et/ou de l'ovaire et/ou du poumon. En effet, les protéines selon l'invention sont sur-exprimées dans les tumeurs et notamment dans le cancer de la prostate (US 5804368, WO 97

10503, CS 84 02898, CS 83 02012, CS 82 08506, CS 82 08215), le cancer du sein (Stoecz et Gould, 1995 ; Simard *et al.*, 1992), le cancer de l'utérus, le cancer de l'ovaire, le cancer du poumon.

Les OBP selon la présente invention peuvent également être
5 utilisées dans des compositions pharmaceutiques, ceci notamment afin de vectoriser certaines drogues.

En effet, les lipocalines sont utilisées par les mammifères pour transporter des molécules hydrophobes au sein de fluides biologiques. Il semble même que ce soit les transporteurs naturels
10 des xénobiotiques *in vivo*. Pour exemple, une surcharge expérimentale en xénobiotiques crée chez le rat mâle des tumeurs au niveau du tube contourné proximal (TCP) du néphron (Borghoff *et al.*, 1990). En effet, les protéines MUP, seulement produites chez le rat mâle, sont réabsorbées par les cellules TCP ; après
15 dégradation lysosomiale, les protéines MUP libèrent les xénobiotiques qu'elles transportaient. Ceux-ci s'accumulent dans ces cellules et les conduisent vers une voie tumorale par mutagenèse. Les lipocalines semblent donc être la structure idéale pour piéger en leur calice une molécule utilisée comme
20 médicament, évitant ainsi une distribution générale de celui-ci. Ainsi, G. Beste *et al.*(1999) et la demande WO 99 16873 décrivent une stratégie de mutagenèse et de criblage permettant d'identifier des variants d'une lipocaline ayant une affinité optimale pour un ligand donné. Il est donc possible de créer une « cage à
25 médicament » hautement spécifique.

La présente invention concerne donc un polypeptide selon l'invention en tant qu'agent de ciblage de composé pharmaceutique. Par agent de ciblage, on entend désigner les polypeptides de l'invention capables d'assurer le transport et la libération du ligand

auxquels ils sont liés au niveau de certains tissus ou cellules cibles qui possèdent à leur surface des récepteurs auxdits polypeptides. Ainsi, l'invention concerne le transport de médicament au sein de la cage d'un polypeptide selon l'invention et du ciblage de cellules
5 grâce à la capacité de l'OBPII à se fixer à un récepteur spécifique.

Il est également dans l'étendue de l'invention de développer des protéines de fusion ou des protéines chimériques comportant la cage correspondant aux lipocalines selon l'invention associées à une protéine permettant un adressage cellulaire spécifique
10 généralement différent de celui naturel des OBPII. La présente invention a donc pour objet un polypeptide caractérisé en ce que ledit polypeptide est exprimé sous la forme d'une protéine de fusion avec une protéine permettant un adressage cellulaire spécifique. Parmi les protéines permettant un adressage cellulaire spécifique, il
15 convient de citer les interleukines, les cytokines, les lymphokines, les chémokines, les facteurs de croissance, les hormones, les anticorps mono- ou polyclonal. Les interleukines, cytokines et lymphokines sont choisies dans un groupe composé de préférence des interleukines Il-1 à Il-20, des interférons α -IFN, β -IFN et γ -IFN.
20 Les facteurs de croissance sont de préférence les facteurs stimulateurs des colonies (colony stimulating factors) G-CSF, GM-CSF, M-CSF et l'érythropoïétine. Les hormones sont choisies de préférence parmi les hormones stéroïdiennes.

Selon un autre mode de réalisation, le polypeptide selon
25 l'invention en tant qu'agent de ciblage de composé pharmaceutique est associé à une molécule permettant un adressage cellulaire spécifique. Parmi, les molécules permettant un adressage cellulaire spécifique, il convient de citer le groupe des stéroïdes, des interleukines, des cytokines, des lymphokines, des interférons, des

facteurs de croissance, des hormones, des anticorps. De préférence, la molécule est un stéroïde. L'association entre le polypeptide selon l'invention et ladite molécule est réalisée soit par une liaison non covalente en utilisant par exemple le système avidine-biotine ou
5 soit par une liaison covalente en utilisant par exemple des agents chimiques pontants.

Parmi les composés pharmaceutiques que peut transporter et cibler le polypeptide selon l'invention, il convient de citer les médicaments et notamment les agents anticancéreux. Les agents
10 anti-cancéreux sont sélectionnés parmi le groupe des agents antiprolifératifs, antinéoplastiques ou cytotoxiques et sont utilisés pour arrêter le développement des cancers et pour induire une régression et/ou une élimination de la masse tumorale. Ces agents anticancéreux sont de préférence des radioisotopes, et de manière
15 encore préférée des radioisotopes émetteurs de rayons gamma tels que l'Iode¹³¹, l'Yttrium⁹⁰, l'Or¹⁹⁹, le Palladium¹⁰⁰, le Cuivre⁶⁷, le Bismuth²¹⁷ et l'Antimoine²¹¹. Les radioisotopes émetteurs de rayons beta et alpha peuvent également être utilisés pour la thérapie. Les agents anticancéreux non isotopiques liés au polypeptide selon
20 l'invention sont multiples et variés; on peut citer: (i) les antimétabolites telles les agents anti-folate, le méthotrexate, (ii) les analogues des purines et des pyrimidines (mercaptopurine, fluorouracile, 5-azacytidine), (iii) les antibiotiques, (iv) les lectines (ricine, abrine) et (iv) les toxines bactériennes (toxine diphtérique);
25 les toxines sont choisies de préférence parmi l'exotoxine A de *Pseudomonas*, la toxine diphtérique, la toxine cholérique, la toxine anthrax de *Bacillus*, la toxine Pertussis, la toxine Shiga de *Shigella*, la toxine apparentée à la toxine Shiga, les toxines d'*Escherichia coli*, la colicine A, la d-endotoxine, l'hémagglutinine d'*Haemophilus A*.

L'invention concerne également une composition pharmaceutique comprenant un composé pharmaceutique lié au moins à un polypeptide selon l'invention et un véhicule pharmaceutiquement acceptable. Il est également dans l'étendue de l'invention d'augmenter la capacité de fixation par mutagenèse de cette même protéine. L'invention concerne une composition pharmaceutique telle que définie précédemment pour le traitement du cancer, et notamment le cancer de la prostate, le cancer du sein, le cancer de l'utérus, le cancer de l'ovaire, le cancer du foie et le carcinome des cellules épithéliales pulmonaires. Etant donné que l'OBP dénommée hOBPIIb, l'apolipoprotéine D, la RBP et l' α -1-acide glycoprotéine sont exprimées dans la prostate, la composition pharmaceutique selon l'invention est destinée au traitement du cancer de la prostate. De même, étant donné qu'il a été démontré que l'ARNm des protéines ApoD, RBP et LCN1 est produit dans la glande mammaire en plus de la production de hOBPIIb déjà décrit dans la demande WO 99 07740, la composition pharmaceutique selon l'invention est destinée au traitement du cancer du sein.

L'invention porte aussi sur un polypeptide selon l'invention en tant que transporteur de composé pharmaceutique. Par transporteur, on entend désigner des polypeptides selon l'invention capable de véhiculer dans l'organisme un composé pharmaceutique sans que ledit composé ne soit libéré à un endroit privilégié dans l'organisme. Un tel polypeptide constitue un moyen de délivrance dudit composé pharmaceutique dans l'organisme.

La présente invention concerne également une composition pharmaceutique selon l'invention caractérisé en ce que ledit polypeptide constitue une forme retard de délivrance dudit composé pharmaceutique dans l'organisme.

En effet le composé pharmaceutique d'intérêt lié au polypeptide OBPII selon l'invention diffuse progressivement dans l'organisme à mesure que ledit polypeptide transporteur est catabolisé dans l'organisme. L'utilisation des techniques de l'ADN recombinant permet à l'homme de l'art de modifier la durée de
5 demi-vie du polypeptide OBPII dans l'organisme en introduisant des modifications dans ledit polypeptide. Ainsi, il peut être intéressant de développer un polypeptide homologue selon l'invention pour lequel les sites de coupures protéasiques ont été mutés afin
10 d'augmenter la demi-vie dudit polypeptide dans l'organisme. Il peut être également intéressant de développer un polypeptide multimérique, exprimé par exemple sous la forme d'une protéine de fusion, afin d'éviter une « fuite glomérulaire » (clearance rénale) et ainsi d'augmenter la demi-vie dudit polypeptide dans l'organisme.

15 Selon un autre aspect, l'invention concerne un composé caractérisé en ce qu'il est choisi parmi un anticorps, un polypeptide, un ligand, un polynucléotide, un oligonucléotide ou un vecteur selon l'invention à titre de médicament et notamment en tant que principes actifs de médicament ; ces composés seront
20 préférentiellement sous forme soluble, associés à un véhicule pharmaceutiquement acceptable. Par véhicule pharmaceutiquement acceptable, on entend désigner tout type de véhicule employé habituellement dans la préparation de compositions injectables, c'est-à-dire un diluant, un agent de
25 suspension tel une solution saline isotonique ou tamponnée. De préférence, ces composés seront administrés par voie systémique, en particulier par voie intraveineuse, par voie intramusculaire, intradermique ou par voie orale. Leurs modes d'administration, posologies et formes galéniques optimaux peuvent être déterminés

selon les critères généralement pris en compte dans l'établissement d'un traitement adapté à un patient comme par exemple l'âge ou le poids corporel du patient, la gravité de son état général, la tolérance au traitement et les effets secondaires constatés, etc.

5 L'invention concerne également une composition pharmaceutique comprenant un vecteur d'expression selon l'invention et un véhicule pharmaceutiquement acceptable.

 Il est également possible d'utiliser la propriété d'expression dans des sites particuliers de ces protéines afin d'assurer une
10 imagerie et/ou un diagnostic de certains types de cancer.

 Enfin, compte tenu de la forte expression des polypeptides selon l'invention dans le placenta (cf. figure 10), la présente invention présente différentes applications thérapeutiques destinées à traiter des pathologies de la gestation chez la femme.

15 Ainsi, un des objets de la présente invention est d'utiliser le polypeptide selon l'invention pour la préparation d'un médicament destiné au transport de composés à travers la barrière placentaire. Ces composés sont de préférence des molécules lipophiles. Le polypeptide selon l'invention peut être utilisé pour réaliser le
20 transport de composés de la mère gestante vers le fœtus ; les composés pouvant être choisis parmi les molécules d'intérêt thérapeutique, de préférence lipophiles, les acides gras essentiels, les hormones, les vitamines, et de manière plus générale, les petites molécules lipophiles.

25 En effet, il a été démontré que les cellules des villosités placentaires sécrètent massivement des polypeptides selon l'invention, notamment des OBPII. Ces polypeptides se répandent ensuite dans le sang fœtal et vers la circulation sanguine générale du fœtus.

Egalement, le polypeptide selon l'invention peut être utilisé pour réaliser le transport de composé du fœtus vers le mère gestante. En effet, le fœtus sécrète au niveau des villosités des structures nasales des polypeptides selon l'invention qui participent aux échanges de molécules entre le fœtus et la mère au niveau de la membrane amniotique. C'est donc également un objet de la présente invention de fournir au fœtus des polypeptides selon l'invention pour permettre le transport de composés présents dans le fœtus, par exemple des xénobiotiques, vers le placenta et/ou vers la membrane amniotique, pour permettre la détoxification du fœtus.

Selon un autre aspect de l'invention, le polypeptide selon l'invention peut être utilisé comme marqueur de grossesse. En effet, le polypeptide selon l'invention est présent dans différents liquides biologiques de la mère tels l'urine, le sang ; la mesure de sa concentration et/ou de sa présence constitue un indicateur de grossesse. A ce titre, ce marqueur pourra être utilisé, avec, ou à la place du, ou en complément du marqueur de grossesse que constitue la choriogonadotrophine humaine hCG ; le polypeptide selon l'invention et la hCG sont en effet des molécules qui sont sécrétées par les cellules placentaires trophoblastiques.

Enfin, selon un autre aspect de l'invention, le polypeptide peut être utilisé comme marqueur de pathologies fœto-placentaires. Il peut constituer un marqueur de la rupture de la membrane de la poche des eaux.

D'autres caractéristiques et avantages de l'invention apparaissent dans la suite de la description avec les exemples et les figures dont les légendes sont représentées ci-après.

FIGURE 1 : Organisation génomique du locus LCN1/hOBPII

La ligne du haut représente la région du chromosome 9q34. La double flèche indique l'intervalle entre la localisation des marqueurs polymorphes D9S1811 et D9S67 ; la position relative des marqueurs D9S67 et D9S1826 est incertaine.

Le niveau du milieu indique l'organisation partielle des cosmides à 3 loci différents : LCN1c, LCN1b-hOBPIIb, LCN1-hOBPIIa. Les cosmides (approximativement 40 kb) sont représentés par des lignes horizontales avec leur nom et les bras des vecteurs T3 ou T7 notés au-dessous de la ligne. Les flèches représentent les différents gènes ou pseudogènes avec leur orientation respective. Dans les cosmides, les lignes en pointillés verticales représentent les sites EcoRI. Le symbole \otimes indique une orientation incertaine du locus.

Le niveau inférieur montre les structures intron/exon des gènes LCN1 et hOBPII : les boîtes noires représentent les exons ; les flèches représentent le site d'initiation de la transcription ; ol1 à ol5 représentent les sondes d'oligonucléotides utilisées pour cribler les cosmides ; *Alu* représente la présence de séquences répétées.

FIGURE 2 : Analyse "Dot Plot" de :

(A) séquence locus LCN1-hOBPIIa (AC000396+ACXXXXXX) contre locus LCN1b-hOBPIIb (AC002098) ;

(B) séquence locus LCN1-hOBPIIa (AC000396+ACXXXXXX)

contre

séquence locus LCN1c ;

(C) séquence locus LCN1b-hOBPIIb (AC002098) contre séquence locus LCN1c (AC002106).

Des séquences génomiques ont été filtrées pour les séquences répétitives à l'aide de « Repeatmasker » (REF), puis comparées et « dotplotted » avec le logiciel gcg en utilisant une taille de fenêtre de 25, et un critère de 20 avec une homologie de 80%.

5

FIGURE 3 : Séquence nucléotidique de gènes hOBPII.

Les lignes supérieures représentent la séquence hOBPIIa et les lignes inférieures la séquence hOBPIIb pour laquelle seuls les nucléotides différents sont représentés ; un tiret indique l'absence de séquences correspondantes. Les capitales ombrées sont les séquences exoniques et les lettres minuscules les séquences introniques. Les tailles indiquées sur la gauche sont indiquées en pb. La boîte TATA est en caractères gras et le signal de polyadénylation est souligné. Les boîtes indiquent les sites accepteurs d'épissage pour les exons 5, 5b, 5c.

15

FIGURE 4 : Représentation schématique des deux gènes hOBPII et de leur ARNm correspondants.

Les lignes horizontales représentent l'organisation exon/intron avec des tailles indiquées en pb. Les boîtes ombrées numérotées de 1 à 7 sont les séquences exoniques codantes des principaux transcrits ; les lettres b et c font référence à des exons surnuméraires. Les différents transcrits sont représentés à l'aide de boîtes assemblées : les premières, hOBPIIa α et hOBPIIb α correspondent aux transcrits principaux, les autres correspondent aux formes issues d'un épissage alternatif. ↓ indique un décalage du cadre de lecture résultant de l'insertion ou de la délétion d'un exon et * représente un codon stop. La lettre « a » représente une

20

25

hélice α et « b » des feuilles β prédits par le logiciel DSC. Les lettres en italique sont des prédictions obtenues avec le logiciel "Predator".

FIGURE 5 : Alignement des séquences des protéines issues des
 5 **deux gènes hOBPIIa et hOBPIIb humains (hOBPIIa α =**
OBP2aaHOMS, hOBPIIb α = OBP2baHOMSA, hOBPIIb β =
OPB2bbHOMSA, hOBPIIa γ = OBP2agHOMSA, hOBPIIa β =
OBP2abHOMSA), des lipocalines des larmes humaines
 10 **(LCN1_HOMSA), d'OBPII de rat (OBP2_RATNO), de**
lactoglobuline bovine BLG (LACB_BOSTA), de MUP de souris
(MUP6_MUSMU), de RBP humaine (RBP_HOMSA), d'OBP bovine
(OBP_BOSTA), de MUP de rat (MUP_RATNO), d'OBP porcine
(OPB_SUSSC).

15 Les résidus dans les boîtes gris foncé sont identiques et
 ceux dans les boîtes gris claires sont similaires. Les éléments de
 structure secondaire prédits avec le programme DSC sont soulignés
 et les résidus d'acides-amino sont en italiques. Les feuillets β et les
 hélices α sont numérotées pour hOBPIIa et b. Le site de clivage
 20 prévu du peptide signal est indiqué par une flèche (AAA↓LS) en
 position 15. Des séquences non alignées de formes très divergentes
 de gènes épissés, hOBPIIa δ (OPB2adHOMSA) et hOBPIIb γ
 (OBP2bgHOMSA) ont été ajoutées au bas de l'alignement après
 l'analyse.

25 **FIGURE 6 : Analyse RT-PCR**

(A) Détection des produits RT-PCR LCN1, LCN1b, LCN1c
 avec leurs sondes spécifiques, et (B) Détection des produits RT-PCR
 hOBPIIa et hOBPIIb avec leurs sondes spécifiques. (C) ARN G3PDH
 contrôle. Les tailles sont indiquées en pb.

FIGURE 7 : Localisation tissulaire des ARNm hOBP

Des sections du méat moyen (A,B), des cornets (C,D), de la prostate (E,F), des canaux déférents (G,H), des glandes mammaires (I, J) ont été hybridées à des ribosondes de hOBP marquées à la digoxigénine-11-UTP. L'hybridation avec une ribsonde hOBP sens n'a pas révélé de signal (B,D,F,H, J). Un signal d'hybridation spécifique est obtenu avec une ribsonde antisens (A,C,E,G, I). Les flèches indiquent des structures différentes : AC : cellules acineuses, EC : cellules épithéliales, GC : cellules glandulaires, SD : canal sécrétoire et L : lumen (X200).

FIGURE 8 : Arbre de distance phylogénétique des lipocalines des vertébrés.

Les abréviations usuelles des lipocalines ont été employées. Après le symbole "_" sont indiquées les trois premières lettres du genre suivies des deux premières lettres de l'espèce.

FIGURE 9 : Schéma de l'évolution de la sous-famille de gène LCN1-OBPII au cours de l'évolution.

Les boîtes indiquent les exons disposés sur une ligne représentant l'ADN génomique. Les « // » entre les lignes indiquent que les loci ne sont pas consécutifs, sans qu'il soit possible de déterminer l'ordre. La large croix sous le symbole « ? » illustre un événement de duplication partielle ou de complète duplication avec une délétion génomique ultérieure pour le locus LCN1c (choix indiqué par le symbole « * »). La croix plus petite indique l'élimination du septième exon qui semble être spécifique de l'homme à cause des nombreuses séquences répétées Alu présentent dans cette région génomique et parce que le rat a deux

gènes VEGP. Le symbole « ** » marque l'étape de recrutement d'exons basée sur les données présentes et celles de la littérature ; cette étape a pu apparaître à n'importe quel moment après les duplications des loci et a pu être séquentielle.

5

FIGURE 10 : Analyse par RT-PCR dans la sphère orale et la sphère génitale de l'expression des autres lipocalines humaines connues.

Définition de nouvelles fonctions OBP et VEGP pour les protéines RBP (retinol-binding-protein) et ApoD (apolipoprotéine D).

10

FIGURE 11 : Etude de l'anticorps polyclonal dirigé contre les protéines hOBPII.

La protéine de fusion est déposée dans les puits 1 et 3. Un échantillon de mucus nasal humain est déposé dans les puits 2 et 4. Les protéines contenues dans les puits 1 et 2 ont été révélées par coloration au bleu de Coomassie. L'anticorps hybridé sur la membrane obtenue par Western-blotting révèle spécifiquement la protéine GST-hOBPIIb entre 30 et 42 kDa (puits 3) et les hOBPIIa et hOBPIIb (puits 4) de 18 kDa dans le mucus nasal humain.

15

20

FIGURE 12 : Immunohistochimie sur les tissus de l'appareil olfactif.

Une forte immunoréactivité est repérable sur la coupe de septum (coloration verte, A) sur les deux coupes de cornet (C et D) et sur la coupe de méat moyen (F). Les témoins négatifs réalisés pour le septum (B) et pour le cornet (E) et pour le méat moyen (G) ne montrent aucune réactivité. Les noyaux des cellules sont repérés par une coloration au DAPI (coloration bleue). Les grossissements

25

utilisés sont de X 100 pour les coupes A, B, C, E et G et de X 200 pour les coupes D et F.

FIGURE 13 : immunohistochimie sur tissus de la sphère orale

5 Ces deux tissus, glandes lacrymales (A, B) et glandes de Von Ebner (E, F) montrent une forte immunoréactivité en comparaison des résultats obtenus grâce au sérum pré-immun (C pour les glandes lacrymales et D pour les glandes de Von Ebner). Grossissement X200.

10

FIGURE 14 : Immunohistochimie sur les glandes mammaires et poumon

15 Une forte immunoréactivité est visible sur les coupes de glandes mammaires (A et B) comparé au témoin négatif (C) réalisé grâce au sérum-pré immun. Une forte immunoréactivité est également observée sur les coupes de poumon (E et F) comparée toujours au témoin négatif (D). chaque observation est faite à un grossissement X 100 pour A et C et à un grossissement X 200 pour B, D, E et F.

20

Figure 15 : Mise en évidence des protéines hOBPII dans différentes structures nasales : Cornet (A : 200x, B : 400x), Septum (C : 300x, D 600x), Méat moyen (E : 400x, F : 600x) :

25 Détection des protéines hOBPII avec un sérum polyclonal de lapin révélé par une activité peroxidasique détecté par DAB. AC : cellule acineuse sécrétrice, P : précipité de DAB (détection hOBPII)

Figure 16 : Mise en évidence de la protéine LCN1 dans différentes structures nasales : Cornet (A : 200x, B : 500x), Septum (C : 100x, D 500x), Méat moyen (E : 200x, F : 600x) :

- 5 Détection de la protéine LCN1 avec un sérum polyclonal de lapin révélé par une activité peroxidasique détecté par DAB. AC : cellules acineuse sécrétrice, P : précipité de DAB (détection LCN1)

Figure 17 : Mise en évidence de la protéine ApoD dans différentes structures nasales : Cornet (A : 500x, B : 700x), Septum (C : 300x), Méat moyen (D : 400x) :

- 10 Détection de la protéine ApoD avec un anticorps monoclonal de souris révélé par une activité peroxidasique détecté par DAB. AC : cellule acineuse sécrétrice, P : précipité de DAB (détection ApoD)

15

Figure 18 : Mise en évidence de la protéine hOBPII dans différentes structures de la sphère orale : Glande Lacrymale (A : 200x, B : 600x), Glandes de Von Ebner (C : 100x, D : 400), Poumon (E : 200x, 800x) :

- 20 Détection de la protéine hOBPII avec un serum polyclonal de lapin révélé par une activité peroxidasique détecté par DAB. AC : cellule acineuse sécrétrice, P : précipité de DAB (détection hOBPII)

Figure 19 : Mise en évidence des protéines hOBPII dans différentes structures placentaires : villosités chorales (A : 300x, B : 800x), Membrane Amniotique (C : 800x), Cordon (D : 600x) :

- 25 Détection des protéines hOBPII avec un serum polyclonal de lapin révélé par une activité peroxidasique détecté par DAB. L : lacune

(contient sang maternel), C : capillaire fœtal, Am : membrane amniotique, E : cellules epitheliales dérivées de l'ectoblaste extra-embryonnaire, MB : membrane basale, M : couche mesenchymateuse, P : précipité de DAB (détection hOBPII)

5

Figure 20 : Mise en évidence de la protéine LCN1 dans différentes structures placentaires : villosités choriales (A : 200x, B : 800x), Membrane Amniotique (C : 600x), Cordon (D : 500x) :

- 10 Détection de la protéine LCN1 avec un serum polyclonal de lapin révélé par une activité peroxidasique détecté par DAB. L : lacune (contient sang maternel), C : capillaire fœtal, Am : membrane amniotique, E : cellules epitheliales dérivées de l'ectoblaste extra-embryonnaire, MB : membrane basale, M : couche
- 15 mesenchymateuse, P : précipité de DAB (détection LCN1)

Figure 21 : Mise en évidence de la protéine ApoD dans différentes structures placentaires : villosités choriales (A : 100x, B : 500x), Membrane Amniotique (C : 400x), Cordon (D : 400x) :

- 20 Détection de la protéine ApoD avec un anticorps monoclonal de souris révélé par une activité peroxidasique détecté par DAB. L : lacune (contient sang maternel), C : capillaire fœtal, Am : membrane amniotique, E : cellules epitheliales dérivées de
- 25 l'ectoblaste extra-embryonnaire, MB : membrane basale, M : couche mesenchymateuse, P : précipité de DAB (détection ApoD)

EXEMPLES

EXEMPLE 1 : MATERIELS ET METHODES

5 **1.A. Banque de cosmides du chromosome 9 humain**

Une copie de la banque de cosmides spécifiques du chromosome 9 humain LL09NC01P construite par le Dr. J. Allmeman (Biochemical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA USA) sous l'égide du National Gene Library Project supporté financièrement par le département américain de l'Energie a été utilisée. Le criblage de la banque et l'analyse des clones ont été réalisés tel que décrit précédemment (Lacazette *et al.*, 1997).

15 **1.B. Clonage et analyse de séquence**

Une banque Lambda gt11 d'ADNc de testicules humains (Clontech) (10^7 p.f.u.) a été amplifiée par 30 cycles de polymérisation en chaîne (PCR) (94°C 45 sec, 54 °C 45 sec, 72°C 1min 30 sec) avec l'amorce oliEST58 CCTGCAGGTACATGAGCTTCC et des amplimères 5' ou 3' de criblage d'inserts situés sur les bras du vecteurs Lambda gt11. Une PCR nichée a ensuite été réalisée avec oliEST26 CGCTGTATTTGCCAGGCTCC et des oligonucléotides spécifiques du bras du vecteur. Les produits PCR ont été sous-clonés dans le vecteur pGEM-T(r), ce qui a permis d'obtenir l'extrémité 5' des ADNc du gène hOBPII.

Les séquences obtenues en utilisant l'oligonucléotide standard pGEM-T (r) et en utilisant un mixe réactionnel prêt à l'emploi de séquençage à base de colorant terminateur (Applied Biosystems) ont été séparées par électrophorèse en utilisant un

séquenceur automatique ABI PRISM 377 (Perkin Elmer) puis ont été analysées avec le logiciel Sequence Navigator 1.0.1. (Perkin Elmer). Des clones d'ADNc pleine longueur de hOBPIIa (hOBPIIa α , hOBPIIa β , hOBPIIa δ , hOBPIIa γ) et hOBPIIb (hOBPIIb α , hOBPIIb β , hOBPIIb γ) ont été obtenus à partir de la RT-PCR par purification des bandes d'intérêt selon les instructions du fabricant (gel extraction kit de Qiagen) ou en sous-clonant dans un vecteur pGEM-T(r) les produits de la PCR nichée pour les formes alternatives faiblement exprimées.

10

1C. Analyse par RT-PCR

Des échantillons de tissus ont été collectés chez des individus caucasiens âgés de 45 à 55 ans en accord avec la réglementation française en vigueur. L'ARN total est extrait selon une méthode en une seule étape utilisant le réactif ARN NOW® selon les instructions du fabricant (Biogentex). 5 μ g d'ARN total ont été rétro-transcrits dans un volume final de 20 μ l en utilisant 0,5 ng d'oligonucléotide GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT avec le système de préamplification Superscript® (Gibco BRL). Trois μ l de ces réactions sont ensuite utilisés pour les PCR suivantes. L'expression des ARNm spécifiques a été déterminée par PCR en utilisant : les amorces TL : CCTCTCCCAGCCCCAGCAAG et AP : GACTCGAGTCGACATCG pour les gènes de type LCN1 (LCN1, LCN1b, LCN1c) et pour les gènes de type hOBPII, les amorces DE : CGCCCAGTGACCTGCCGAGGTC, et FI : CTTTATTTGGAGTCAGGTGGGTG. Comme contrôle de qualité des ARN, les amorces G3PDH1 : CTCTGCCCCCTCTGCTGATG et G3PDH2 : CCTGCTTCACCACCTTCTTG du gène G3PDH ont été

15

20

25

utilisées ; le gène G3PDH est considéré comme étant constitutivement exprimé dans tous les types cellulaires. 32 cycles de PCR (94° C 45 sec., 54° C 45 sec., 72° C 2 min. 30 sec.) ont été réalisés et les produits d'amplification ont été séparés sur gel d'agarose à 1%. L'ADN est transféré sur une membrane Hybond N+®.

Pour la détection de l'expression des différents gènes, plusieurs oligonucléotides spécifiques des gènes respectifs ont été synthétisés :

- 10 - oLCN1 : GACTCAGACTCCGGAGATGA,
- oLCN1b : AACTCAGACACCAGAGATGA,
- oLCN1c : GACTCAGATCCCGGAGATGA,
- ol5 : CCAGGAGGGACCACTACA spécifique du gène hOBPIIb,
- ol4 : CCGGGACGGACGACTACG spécifique du gène hOBPIIa,
- 15 - G3PDH3 : CTCATGACĈACAGTCCATGC,

Les oligonucléotides sont marqués au $\gamma^{32}\text{P}$ -ATP en utilisant de la T4 kinase (Applied Biosystem) ; les hybridations des oligonucléotides marqués sont réalisées à 42° C. Le lavage final est réalisé dans une solution de 0,1 X SSC, 0,1% SDS à 48° C pendant 20 min. La spécificité des réactions d'hybridation des oligonucléotides est contrôlée en utilisant des échantillons d'ADN cosmétique digéré (p233G2 pour LCN1 et hOBPIIa, P19E7 pour LCN1b et hOBPIIb, P181A9 pour LCN1c) et chargés sur le gel avec des produits de RT-PCR.

25

1.D. Génotypage et analyse de liaison

Le génotypage est réalisé par des réactions PCR en utilisant 100 ng d'ADN génomique provenant de 8 familles de référence du CEPH et utilisant les oligonucléotides oli9

TGTTCGGGAACGCAGCTT et oli10bis
TGCCGCTGTCCCCACGTCGG. Les paramètres du thermocycleur
consistent en un cycle initial à 94° C pendant 10 min suivi par 30
cycles à 94° C pendant 30 s, 55° C pendant 30 s et 70° C pendant
5 45 s puis en une étape d'élongation finale de 10 min à 70° C. Les
produits PCR sont ensuite analysés sur un gel d'agarose à 3%. Les
informations concernant les marqueurs du chromosome 9 peuvent
être obtenues à l'adresse Internet suivante ([hTTP://galton.ucl.ac.uk](http://galton.ucl.ac.uk)) ;
les analyses ont été réalisées en utilisant les outils d'études de
10 liaison préalablement décrites dans Lacazette *et al.* (1997). Les
haplotypes sont reconstruits manuellement selon les événements
de recombinaison préalablement décrits dans la famille 1362
(Attwood *et al.*, 1994).

15 **1.E. Prédiction de structures secondaires**

Un alignement multiple des protéines lipocalines pour
lesquelles les structures cristallographiques ont déjà été décrites
(Monaco *et al.*, 1992 ; Spinelli *et al.*, 1998) avec les protéines
hOBPIIa et hOBPIIb a été obtenu en utilisant le logiciel Clustalx
20 (<ftp.infobiogen.fr>). Les structures secondaires putatives ont été
déterminées avec le programme DSC (Discrimination of protein
Secondary structure Class) développé par R.D. King et M.J.E.
Sternberg (<http://bioweb.pasteur.fr/seqanal/interfaces/dsc-simple.html>). Les
structures secondaires des protéines correspondantes aux formes
25 alternativement épissées sont supposées identiques aux formes
traditionnelles avant le décalage du cadre de lecture; après ce
décalage la prédiction de structure est effectuée avec une seule
séquence et est réalisée en utilisant le logiciel Predator
(<http://pbil.ibcp.fr/cgi-bin>).

1.F. L'analyse phylogénétique

Les protéines lipocalines dont les séquences sont connues dans leur totalité ont été alignées trois fois consécutivement en utilisant le logiciel Clustalx (<ftp://ftp.inforbiogen.fr>). Des distances dans l'arbre phylogénétique ont été calculées avec Clustalx et dessinées avec Njplot.

1.G. Hybridation in situ

Des coupes sériées réalisées au cryostat (8 μ m d'épaisseur) sont collectées sur des lames SuperFrost® Plus (Menzel Glazer) et stockées à - 80° C. Des sondes ARN antisens et sens sont obtenues selon les techniques standard en utilisant la T7 ou la SP6 polymérase à partir d'une matrice obtenue par digestion avec des enzymes de restriction NcoI ou PstI du clone cDNA de phOBPIIaP2 et en utilisant du DIG-11-UTP (Boehringer Mannheim) (la longueur de la sonde est approximativement de 150 nucléotides). Les matrices digérées par PstI et transcrites par l'ARN polymérase de T7 correspondent à la sonde antisens et les matrices digérées par NcoI transcrites par l'ARN polymérase SP6 correspondent à la sonde sens. Des sections de tissus sont fixées dans du paraformaldéhyde 4% pendant 15 min puis rincées pendant 5 min dans du PBS 2X. Après acétylation (2 x 5 min dans un tampon de triéthanolamine (TEA) pH 8, contenant 0,25% v/v d'anhydride acétique), les sections tissulaires sont préhybridées à 60° C pendant 15 min dans de la formamide 50%/1X SSC. Les sondes marquées sont appliquées sur chacune des sections dans 50 μ l de tampon d'hybridation (50% formamide, 1X Denhardt's, 500 μ g/ml total d'ARNt, 10% de Dextran sulfate, 10 mM de dithiotréitol). Les sections sont recouvertes puis incubées dans des chambres humides à 50° C

pendant la nuit. Après hybridation, les lames sont immergées à 55° C dans un tampon de lavage (50% formamide, 1X SSC) pendant 2 heures. Chaque lame est ensuite rincée 2 fois 5 min dans du 2X SSC à température ambiante, puis traitée pendant 30 min avec 10 mg/ml de RNase à 37° C, et enfin immergée 2 heures à 55° C dans une solution de lavage (50% formamide, 2X SSC). Les lames sont ensuite placées pendant 15 min à 55° C dans du 0,1 X SSC. La détection immunologique est réalisée en utilisant un anticorps anti-DIG conjugué à la phosphatase alcaline (fragments Fab) selon le protocole de Boehringer Mannheim. Les sections sont examinées à des grossissements différents utilisant un microscope Axiophot (Zeiss).

1.H. Obtention d'un anticorps polyclonal dirigé contre les protéines hOBP_{II}

Deux cent cinquante µl de protéines de fusion ont été fournis à la société Agro-Bio pour la production d'anticorps. Des lapins (New Zealand White SPF) sont injectés aux jours 0, 14, 28 et 42. Les serums sont prélevés aux jours 0, 35, 49 et 63. Afin de tester l'anticorps, la protéine de fusion est déposée dans les pistes 1 et 3 (Figure 11), alors que du mucus nasal humain est déposé dans les pistes 2 et 4. Les pistes 1 et 2 ainsi que la piste correspondant à l'échelle de poids moléculaire, correspondent à une révélation par le bleu de Coomassie. Les pistes 3 et 4 correspondent à une détection des protéines reconnues par l'anticorps à l'aide de la technique de Western blot. Les pistes 1 et 3 montrent la présence d'une série de protéines recombinantes tronquées entre 30 et 42 kDa issues d'une synthèse peu efficace due à la présence de nombreux codons rares chez la bactérie au sein de la séquence de hOBPIIb. Toutefois cette

production de protéine recombinante fût suffisante pour produire un anticorps polyclonal de bonne qualité comme indiqué par la révélation dans la piste 4 correspondant au mucus nasal, d'une bande de 18 kDa spécifique des protéines hOBPII.

5

1.I. immunohistochimie

1.I.1. Des coupes sériées de 8 μ m de différents tissus sont fixées au para-formaldéhyde (5min), et rincées trois fois au PBS x1 (15 min) puis incubées 30 min dans une solution 3% de BSA en PBSx1, avant d'être incubées durant la nuit en présence de l'anticorps anti-protéine de fusion en solution PBSx1. Après trois rinçage en PBSx1 (15 min), un anticorps anti-IgG de lapin couplé au FITC (coloration verte en fluorescence) en solution de PBSx1 est placé 3 heures au contact des lames. Après un rinçage en PBSx1, une solution de DAPI (100 ng / μ l en PBSx1) est appliquée pendant 10 min (contre coloration des noyaux cellulaires en bleu). Après trois rinçages en PBSx1 (15 min), les lames sont montées dans de l'eau glycinée (50/50). L'analyse est réalisée en présence de filtres DAPI et FITC à l'aide d'une caméra CDD en utilisant un temps d'intégration entre 4 et 32 ms.

1.I.2. Les tissus sont congelés à -80°C en présence d'OCT. Des coupes sériées de 7micromètres sont réalisées à l'aide d'un microtome. Les tissus sont ensuite fixés à la Para-Formaldéhyde 4% en PBS. Les lames sont incubées pendant 10 min dans une solution de méthanol en présence de 4,5% de H_2O_2 , puis rincées trois fois en PBS pendant 5 min. Après une incubation pendant 30 min dans une solution de 3% d'albumine de sérum de bœuf, le premier anticorps (dilué au 1/350) est incubé une nuit à

température ambiante (Anticorps polyclonaux de lapin pour hOBPII et LCN1, et monoclonal de souris pour ApoD). Puis, après trois rinçages de 15 min en PBS, le deuxième anticorps (dilué au 1/300) (anti-lapin pour hOBPII et LCN1, et anti-souris pour ApoD) est
5 incubé 3 h à température ambiante. Après trois rinçages en PBS pendant 15 min, une révélation à la peroxydase par le DAB est réalisé pendant 2 à 10 min, à l'aide du kit « Vector ». Les lames sont ensuite rincées à l'eau, et contre colorées à l'hématoxyline de Mayer pendant 2 min. Après deux rinçages à l'eau, les lames sont
10 déshydratées par bassinage successif dans des solutions d'éthanol 70°, 95°, 100°, 100°, et finalement deux bains de toluène. Les lames sont ensuite montées en baume du Canada, et observées à l'aide d'un microscope Zeiss.

15

**EXEMPLE 2 : IDENTIFICATION D'UN GENE HOMOLOGUE A
LCN1 LOCALISE SUR LE CHROMOSOME 9
HUMAIN**

L'identification de l'ADN complémentaire de LCN1 codant
20 pour la lipocaline des larmes humaines (Lassagne et Gachon, 1993) a préalablement été rapportée, ainsi que la localisation du gène sur le chromosome 9q34 (Lassagne *et al*, 1993 ; Lacazette *et al*, 1997). Les deux gènes codant pour les protéines des glandes de von Ebner 1 et 2 (Kock *et al*, 1994) correspondent aux protéines homologues
25 de LCN1 chez le rat ; ceci pose la question de savoir s'il existe d'autres gènes codant pour une protéine LCN1 dans le génome humain. Les expériences d'hybridation *in situ* (Lassagne *et al*, 1993) ainsi que les analyses avec des hybrides somatiques (Lassagne *et al*, 1995) indiquent que, s'ils existent, ces gènes humains

additionnels doivent se trouver localisés dans la région chromosomique 9q34.

Une banque de cosmides spécifiques du chromosome 9 humain (LL09NP01) a été criblée avec la sonde d'ADNc LCN1 humaine ; 26 cosmides ont été identifiés ; ceux-ci ont été digérés par EcoRI ou PvuII, puis hybridés successivement avec l'ADN de LCN1 et différents oligonucléotides (figure 1).

Les cosmides sont divisés en 3 groupes. Le premier groupe (clones P32H3, P41B5, P63B6, P92H20, P109C6, P145H6, P195B4, P233G2, P233F2, P265D4, et P276H8) correspond aux cosmides contenant une séquence du gène LCN1 préalablement isolé (numéro d'accension : L14927) formé par 7 exons (Holzfeind et Redl, 1994). Le second groupe (clone P19E4, P19E7, P42H9, P98H5 et P142H8) correspond à la séquence LCN1b homologue à LCN1 (clone P19E4) (numéro d'accension : Y10826) depuis le promoteur jusqu'au 6^{ème} exon puis qui diverge ensuite. Un troisième groupe de cosmides (clone P110C1, P174E4, P174E5, P181A9, P181B6, P211A7, P238C6 et P291E1) contient une région LCN1c, établi à partir du séquençage partiel du clone P181A9 (numéro d'accension Y10827), qui est fortement homologue à LCN1 seulement à partir du promoteur jusqu'à l'exon 2. Ainsi, LCN1 est le seul gène qui possède le 7^{ème} et dernier exon. De plus, les boîtes TATA sont dégénérées dans les promoteurs des gènes LCN1b et LCN1c.

25

**EXEMPLE 3 : IDENTIFICATION DE DUPLICATIONS
GENOMIQUES CONTENANT LES GENES DE
LIPOCALINES ET CARTOGRAPHIE DE LA
REGION CHROMOSOMIQUE 9Q34**

Au cours de la période d'identification de la famille de gène LCN1, un vaste projet de cartographie physique a conduit à l'identification de contigs de cosmides couvrant partiellement la région chromosomique 9q34 (Nahmias *et al.*, 1994, Van Slegtenhorst *et al.*, 1995, Hornigold *et al.*, 1997) ; le séquençage des clones correspondants a été réalisé par un groupe du Massachusset Institute of Technology (Boston, USA).

Des analyses de comparaison de séquences dans les banques de données de LCN1, LCN1b et LCN1c ont révélé des homologues fortes entre le gène LCN1 préalablement décrit (numéro d'accension L14927) et les trois séquences cosmides : P161A1 ayant le numéro d'accension AC002098, P203H12 avec le numéro d'accension AC000396 et P161G2 avec le numéro d'accension AC002106.

Une analyse très fine notamment de la région 3' des gènes LCN1 a révélé que la séquence LCN1c correspondait au clone ayant le numéro d'accension AC002106, LCN1b au numéro AC002098 et LCN1 au numéro AC000396 ; il existe cependant une insertion de 60 paires de base à la position 12360 du clone ayant le numéro d'accension AC000396 par rapport à celui ayant le numéro L14927. Des analyses complémentaires ont révélé que ces séquences sont homologues sur une région beaucoup plus large que les gènes LCN1. Comme illustré par les analyses en dot plot (figure 2), ces trois cosmides correspondent à des régions génomiques dupliquées ; les cosmides P161A1 (AC002098) et P203H12 (AC000396) sont homologues sur la totalité de leurs séquences ; le cosmide P161G2 (AC002106) n'est homologue à la séquence uniquement pour des séquences situés en amont de l'intron 3 du gène LCN1.

Les positions relatives de ces régions dupliquées du chromosome 9 humain ont été analysées. Les résultats ont été comparés avec la carte physique établie dans le cadre du projet de cartographie physique par des analyses de restriction des cosmides (Hornigold *et al*, 1997) en utilisant la même banque LLN09NC01P. Approximativement les mêmes cosmides ont été trouvés dans les deux recherches pour le locus LCN1c ; la comparaison de séquence du cosmide P181A9 révèle que son extrémité T3 contient le gène Surf 5 (figure 1). De plus, AC002106 est localisé dans un contig de séquence entre ABO et le locus Surfeit confirmant la localisation de LCN1c. P161A1 et P203H12 ont des localisations moins précises en utilisant les données de Hornigold et de ses collaborateurs à cause de la limite de la stratégie de cartographie de restriction pour les régions dupliquées. P203H12 (AC000396) semble correspondre au gène LCN1 qui a été préalablement localisé près du marqueur D9S1826 (Lacazette *et al*. 1997) à l'exception d'une insertion de 60 paires de bases.

Pour tester si P203H12 peut correspondre à une 4^{ème} région dupliquée, la région a été testée par PCR sur de l'ADN génomique de 150 individus non apparentés en utilisant des amorces spécifiques pour AC000396 et L14927. Une bande unique correspondant à la longueur de la séquence AC000396 a été détectée (donnée non montrée) prouvant ainsi que P203H12 contient le gène LCN1 préalablement décrit. Le locus LCN1b n'a pas été positionné précisément dans la région comme il est démontré par le fait que les extrémités AC002098 ainsi que les extrémités des cosmides contenant LCN1b ne détectent aucune séquence homologue dans les banques de données.

L'analyse des séquences a permis d'identifier, un mini-satellite putatif à la position 3177-3724 du cosmide correspondant à AC002098 (Figure 1) ; un polymorphisme rare a ainsi été mis en évidence dans une population de 20 individus non apparentés (PIC = 0,05). Ce nouveau marqueur polymorphe n'est informatif que dans la seule famille 1362 sur les 8 familles de référence du CEPH testées.

L'analyse de liaison a révélé des lod-scores à deux points supérieurs à 3 à teta = 0 pour les marqueurs D9S275 et D9S1818. La reconstruction de l'haplotype a confirmé la localisation du gène LCN1b entre les marqueurs D9S1811 et D9S67 dans le chromosome 9q34 (Figure 1).

EXEMPLE 4 : IDENTIFICATION DE DEUX NOUVEAUX GENES DE LIPOCALINE

Les analyses de séquences ont révélé des données nouvelles.

La comparaison de AC002098 à la banque de données a révélé des similarités entre cette séquence et les gènes de lipocalines. En plus de la région 21000 à 27000 qui contient le gène LCN1b, la région autour de la position de 2150 contient des séquences présentant des similitudes avec les séquences codant pour les protéines de liaison aux odeurs de rat de type II, des lipocalines des larmes ainsi que pour l'EST AA460385.

Parallèlement, la comparaison de la séquence AC000396 à la banque de données a révélé en plus de la région 11100 à 17100 contenant le gène LCN1, des similitudes pour le même groupe de séquences dans la région 36600 à 37800.

L'EST AA460385 exprimée dans les testicules humains correspond à 4 exons du nouveau gène de lipocaline présent dans le clone P161A1 (AC002098). Un exon putatif de 50 paires de bases similaire à la séquence EST est également présent à l'extrémité du
5 clone P203H12 (AC000396).

Les inventeurs ont donc conclu à l'existence d'un nouveau gène de lipocaline à une position distale de LCN1b. Les inventeurs ont émis l'hypothèse, que suite à une duplication génomique, un second gène orthologue au gène codant pour l'EST est présent dans
10 la région 3' de LCN1.

Le séquençage du cosmide P233G2 contenant cette région (Figure 1) avec des oligonucléotides définis à partir d'EST a en effet révélé la présence d'un autre gène de lipocaline à un locus distant de 20 kb distal du gène de LCN1. Pour identifier les premiers exons
15 des deux nouveaux gènes, des PCR nichées sur des clones cDNA provenant d'une banque de testicules ont été réalisées entre les oligonucléotides localisés dans la région 5' de l'EST et les bras du vecteur. Les produits PCR sont clonés et leur séquence a révélé trois exons additionnels d'après la séquence génomique (Figure 1).
20 Une boîte TATA est présente en amont du premier exon dans les deux cas (Figure 3). Les deux ARNm, hOBPIIa correspondant au gène localisé en aval de LCN1 et hOBPIIb localisé en aval de LCN1b (Figures 3 et 4), sont identiques à 97,5% et 63 % à ceux correspondant au gène LCN1. Les organisations intron/exon de ces
25 deux gènes sont consistantes avec la famille des lipocalines.

De plus, les séquences protéiques déduites (170 amino acides) confirment l'appartenance à la famille des lipocalines. Les deux protéines matures putatives hOBPIIa (MW = 17,8 kDa) et hOBPIIb (MW = 18,0 kDa) sont identiques à 89%. Ces deux

nouvelles protéines possèdent un peptide signal putatif de 15 amino acides (Figure 5), les résidus conservés GXWY à la position 27 à 30 et deux cystéines susceptibles d'être impliquées dans un pont disulfure (position 74 à 166).

5 Cependant, les comparaisons de séquences par paire avec les autres lipocalines indiquent une faible conservation des résidus acides aminés (15 à 25%) à l'exception de TL/VEG humain codé par le gène LCN1 (valeur moyenne 43 % sur 155 amonoacides) et l'OBPII de rat (valeur moyenne 45,5 %). Par ailleurs, les points
10 isoélectriques calculés de hOBPIIa et de hOBPIIb sont respectivement de 7,85 et de 8,72 alors que ceux des lipocalines sont acides (aux environs de 4,5) à l'exception de l'OBPII de rat (PI = 9,01). La prédiction des structures secondaires des protéines hOBPIIa et hOBPIIb avec le logiciel DSC en effectuant des
15 alignements multiples avec les séquences de lipocalines de structures connues indique la présence de 8 brins de feuillet β antiparallèles pouvant permettre la formation d'un calice suivi par une hélice α et un feuillet β final en accord avec les données connues des structures des lipocalines (Figure 5).

20

EXEMPLE 5 : ETUDE DE L'EXPRESSION DES GENES DE LIPOCALINES IDENTIFIES

Pour préciser si LCN1b, LCN1c, hOBPIIa et hOBPIIb sont
25 des gènes exprimés, des analyses de RT-PCR ont été réalisées sur 18 tissus humains différents connus pour produire des lipocalines (Figure 6).

Deux couples d'amorces ont été synthétisés ; l'un reconnaît les ARNm de type LCN1 et l'autre les ARNm de type hOBPII.

L'hybridation des oligonucléotides spécifiques des gènes sur des membranes sur lesquelles sont immobilisés les produits de RT-PCR après transfert par la technique de Southern, puis le sous-clonage des produits hybridés ont révélés que LCN1b et LCN1c ne sont
5 exprimés dans aucun des produits testés, alors que les ARNm de LCN1 sont détectés dans la glande lacrymale, les glandes sudoripares, les glandes de von Ebner, le septum nasal, l'épithélium du cornet nasal, tout comme le placenta et les glandes mammaires (Figure 6A). Ces données et les informations de
10 séquences qui indiquent que la boîte TATA et le dernier exon ont été perdus permettent d'affirmer que LCN1b et LCN1c sont des pseudogènes qui ne participent pas à la formation des protéines LCN1 humaines.

A l'inverse, les gènes hOBPIIa et hOBPIIb sont exprimés, ce
15 qui confirme leur détection précédente dans les banques d'ADNc. De manière surprenante, bien que les deux protéines hOBPIIa et hOBPIIb soient très similaires sur la totalité de leur séquence y compris dans la région promotrice de 1,5 kb, leur profil d'expression sont différents (Figure 6B). La protéine hOBPIIa est
20 fortement exprimée dans le septum nasal, le méat moyen, le cornet nasal, les testicules et le placenta et plus faiblement dans les glandes mammaires, les glandes lacrymales, les glandes sudoripares, les glandes de von Ebner et le poumon. A l'inverse, la protéine hOBPIIb est exprimée de manière prédominante dans la
25 prostate, les testicules et les glandes mammaires et plus faiblement dans les glandes sous-maxillaires, le septum nasal et le méat moyen.

De plus, les analyses de RT-PCR ont révélé l'existence d'un épissage alternatif du produit de transcription du gène

hOBPIIa et du gène hOBPIIb qui génère respectivement quatre et trois ARNm (Figures 3, 4, 5 et 6).

La transcription du gène hOBPIIa génère au moins quatre ARNm qui codent pour quatre protéines différentes; le premier
5 ARNm code pour la protéine hOBPIIa α qui correspond à la protéine hOBPIIa décrite précédemment. Dans le gène hOBPIIa, trois sites accepteur d'épissage différents ont été identifiés pour l'exon 5 (figure 3 et 4) formant ainsi deux autres variants d'épissage. Un premier variant d'épissage présente un site accepteur pour l'exon 5
10 localisé 49 pb avant le précédent (exon 5b) ; ceci génère un ARNm de 725 nucléotides qui code pour la protéine hOBPIIa β de 146 aminoacides. Cette protéine est identique jusqu'au 8^{ème} feuillet β putatif puis différente ensuite avec seulement 16 aminoacides additionnels. Un second variant d'épissage présente un site
15 accepteur pour l'exon 5 localisé 65 pb avant le précédent (exon 5c) ; ceci génère un ARNm de 741 nucléotides qui code pour une protéine hOBPIIa γ de 228 aminoacides. Cette protéine possède les huit premiers feuillets β putatifs identiques à ceux de hOBPIIa α puis est différente dans la région C-terminale (figure 5) à cause
20 d'un décalage de cadre de lecture généré par cet événement d'épissage alternatif; la structure de cette région C-terminale prédite par le logiciel Predator est une longue région coudée contenant une 9^{ème} feuillet β .

Dans le cas du gène hOBPIIb, en plus de l'ARNm hOBPIIb α
25 préalablement décrit, un exon surnuméraire de 106 pb (exon 3b) entre les précédents exons 3 et 4 a été identifié (figure 3). Cet ARNm plus long (782 nucléotides) code pour une protéine hOBPIIb β de 165 amino-acides. Du point de vue de la structure protéique, hOBPIIb β est identique à hOBPIIb α jusqu'au 5^{ème} feuillet β putatif

puis diffère ensuite à cause d'un décalage du cadre de lecture. Les prédictions des logiciels informatiques indiquent que le motif ALWEALAIIDTRLK est une hélice α qui est juste derrière le cinquième feuillet β . Deux feuillets β additionnels peuvent être
5 présents dans la longue partie C-terminale.

Aucun des variants d'épissage alternatifs pour un gène n'est détecté symétriquement pour l'autre gène bien que les sites d'épissage accepteur et donneur putatifs puissent être présents (figure 3). En plus de ces motifs protéiques qui conservent la
10 structure traditionnelle d'une lipocaline, nous avons démontré, pour les deux gènes hOBPIIa et hOBPIIb, l'existence d'une faible quantité d'ARNm ayant subi un épissage alternatif et codant pour des protéines dont la séquence n'est pas directement liée aux lipocalines. Des ARNm codant pour hOBPIIa δ et hOBPIIb γ
15 auxquels ils manquent la séquence codant pour l'exon 2 et qui possèdent respectivement l'exon 5b et l'exon 5 codent pour des protéines putatives sécrétées de 147 et de 85 amino-acides respectivement (figures 4 et 5); ces protéines divergent des protéines précédentes à partir du 24^{ème} amino-acide.

20 Pour identifier les cellules exprimant les gènes hOBPII, les inventeurs ont hybridé des sondes ribonucléiques sens et anti-sens marquées à la digoxygénine sur des sections tissulaires (figure 7). Les ARNm hOBPII sont détectés dans les cellules acineuses du méat moyen et des cornets nasaux ainsi que des cellules
25 épithéliales des cornets; ceci soutient l'idée que les protéines hOBPII sont impliquées dans la fonction olfactive. En plus de la production dans la sphère orale, des ARNm codant pour hOBPII ont été détectés dans la sphère génitale, notamment dans les cellules glandulaires de la prostate, dans les cellules sécrétrices épithéliales

du canal déférent. Aucun signal n'a été détecté dans les gonades mâles, ce qui suggère que l'expression des gènes hOBPII mise en évidence dans les expériences de RT-PCR correspondait à la présence de canaux additionnels dans la préparation tissulaire (rete testis et canaux efférents). En combinant ces résultats de la détection de l'ensemble des ARNm codant pour hOBPII avec ceux de l'approche RT-PCR, il est apparu que cinq protéines hOBPII (hOBPII α , hOBPII β , hOBPII γ , hOBPII α , hOBPII β) sont sécrétées par les cellules épithéliales des canaux des gonades mâles, ainsi que des cellules acineuses du méat moyen et des cornets nasaux ; dans ces cellules les ARNm codant pour hOBPIIa sont hautement prédominant. Seules les deux protéines hOBPIIb (hOBPII α et hOBPII β) sont sécrétées par les cellules épithélio-glandulaires de la prostate et des glandes mammaires.

EXEMPLE 6 : ANALYSE PHYLOGENETIQUE ET CLASSIFICATION DES LIPOCALINES

L'analyse en dotplot (figure 2) ainsi que les duplications génomiques que nous avons révélées indiquent une origine commune des gènes LCN1 et hOBPII. Pour clarifier les relations entre les membres de la famille des lipocalines et pour vérifier que nous avons identifié le gène humain orthologue de OBPII de rat (Dear *et al.*, 1991), nous avons construit un arbre de distance phylogénétique avec les lipocalines des vertébrés (figure 8).

Les inventeurs ont mis en évidence neuf groupes principaux de lipocalines issus d'un précurseur commun :

- la famille de l'apolipoprotéine et de la protéine de liaison au rétinol (RBP) (groupe 1) ;

- le groupe de la prostaglandine D-synthase et du précurseur de lipocaline associé à la gélatinase des neutrophiles (groupe 2) ;
- la sous-famille de l'alpha-1-microglobuline/bikunin (protéine HC) (groupe 3) ;
- la sous-famille orosomucoïde (A1AG, A1AH, A1AI) (groupe 4)
- la sous-famille de la sphère orale 1 (OBPII-type-LCN1/VEGP, VNSP I et II, LALP, CanF1) (groupe 5) ;
- la sous-famille de la lactoglobuline (groupe 6) ;
- la sous-famille des protéines sécrétées par l'épididyme de lézard (groupe 7) ;
- la sous-famille la sphère orale 2 (protéines urinaires majeures (MUP) de souris) (groupe 8) ;
- la sous-famille de la sphère orale 3 (OBP1, OBPII de souris, Aphrodisine, Probasine, BD20) (groupe 9).

La présente invention concerne plus particulièrement les groupes 5, 8 et 9.

Le groupe 5 contient les protéines hOBPII qui sont étroitement liées d'un point de vue évolutif à l'OBPII de rat. Pris dans leur ensemble, les présents résultats indiquent que les gènes hOBPII humains sont orthologues du gène OBPII de rat. Ce groupe contient également les protéines LCN1-VEGP de différentes espèces. En prenant en considération l'organisation génomique des gènes hOBPII-LCN1, les données de l'arbre illustrent l'événement de duplication (flèche) qui donne naissance aux gènes ancestraux hOBPII et LCN1 à partir de leur précurseur lipocaline commun (Figure 9). Plus récemment, les duplications originelles de la région de 50 kb contenant hOBPII-LCN1 (flèche) ont généré chez l'homme les deux gènes hOBPII et le gène LCN1 et son pseudogène LCN1b.

La duplication additionnelle ayant donné lieu au pseudogène LCN1c est partielle dans le génome humain et ne produit pas de protéine fonctionnelle et est donc absente du présent arbre. De plus, les deux protéines VEG du rat sont plus étroitement liées l'une à l'autre dans l'arbre phylogénétique qu'avec la protéine humaine LCN1. La situation est identique pour les deux protéines humaines OBPII par rapport à la protéine OBPII de rat. Ces résultats sont en faveur d'un processus de conversion génique pour au moins certains gènes des lipocalines. Ceci peut également être
10 corrélé au fait que ces gènes se situent sur le même bras chromosomique.

Le groupe 9 correspond à la famille de la sphère orale 3 et contient des OBP qui ont déjà été décrites. La protéine Aphrodisine a déjà été décrite comme un transporteur de phéromone (Henzel *et al.*, 1988) et apparaît être orthologue à la protéine OBP1 de rat et de souris, avec deux gènes OBP1 paralogues. Finalement, les lipocalines olfactives produites par les glandes de Bowman de *rana pipens* (OLFA RANPI) qui sont considérées comme des transporteurs d'odeurs potentiels dans le mucus de la grenouille,
15 ne sont liées à aucun groupe d'OBP putatif (groupes 5, 8 et 9) suggérant ainsi l'existence d'autres catégories d'OBP.

Certaines lipocalines ont été décrites comme allergènes. L'allergène majeur du chien, la protéine CanF1 majoritairement exprimée dans les glandes de von Ebner (Konieczny *et al.*, 1997),
25 apparaît dans l'arbre être la protéine VEG du chien. Ce résultat montrant que LCN1 peut être allergénique, pousse les inventeurs à proposer l'implication des protéines OBPII dans les processus allergiques. De la même manière, l'allergène majeur du cheval EquC1 exprimé dans le foie, dans les glandes salivaires (Grégoire *et*

al., 1996) apparaît être orthologue à un membre de la protéine MUP (groupe 3). L'allergène majeure de la vache (BD20 BOSTA) qui est présent dans la famille de la sphère orale 2 (groupe 4) est davantage lié à la probasine. CanF2 principalement exprimée dans
5 la glande parotide (Konieczny *et al.*, 1997) ne semble pas être orthologue à une lipocaline préalablement décrite.

La présente invention a permis de révéler l'existence d'une duplication génomique au locus q34 du chromosome 9 humain qui recèle une famille de gène du type LCN1 ; cette famille comporte
10 outre le gène LCN1, précédemment décrit, deux pseudogènes ainsi que deux gènes hOBPII qui sont paralogues à LCN1. Les inventeurs ont révélé que la famille hOBPII-LCN1 résulte d'événements consécutifs de duplications génomiques. Les séquences ainsi que l'organisation génomique ont révélé que les gènes LCN1 et hOBPII
15 dérivent d'un ancêtre commun et ont été générés au moyen de duplication en tandem. Les comparaisons de séquences ont démontré que les protéines hOBPII α et hOBPII β sont les formes de protéines des hOBPII les plus proches de LCN1; ceci est corroboré par le fait que la taille de l'exon 5 de LCN1 est de 121 pb,
20 ce qui correspond à la taille de l'exon 5a de la protéine hOBPII et par le fait qu'aucun exon 3b n'ait été trouvé dans la protéine LCN1. Des résultats similaires sont obtenus lorsque cette comparaison est effectuée avec les sept autres exons des lipocalines. Ces données sont en faveur d'une acquisition de la diversité des protéines
25 hOBPII à travers l'intégration de d'ADN génomique additionnel environnant au niveau du site accepteur d'épissage amont pour l'exon 5 de hOBPIIa ou à travers le recrutement d'un exon surnuméraire (exon 3b) pour la protéine hOBPIIb. L'hypothèse d'un

recrutement d'exons plutôt qu'une réduction de la taille des exons ou qu'une perte d'exons est plus probable.

Les inventeurs ont montré que les gènes hOBPII et LCN1 codent pour des protéines impliquées dans différentes fonctions, 5 comme le montre l'expression de ces gènes à la fois dans la sphère orale et dans la sphère génitale. De plus, les inventeurs ont montré par l'analyse phylogénétique (exemple 6) que plusieurs protéines différentes pouvaient participer à la même fonction d'odorant-binding protéine ; ainsi trois sous-familles de lipocalines (les 10 groupes 5, 8 et 9) correspondant à plusieurs protéines sont retrouvées exprimées dans la sphère orale et notamment dans les glandes nasales et buccales.

Cette analyse montrant que plusieurs lipocalines participent à une même fonction physiologique et qu'une même protéine 15 pouvait participer à différentes fonctions, a conduit les inventeurs à analyser l'expression des autres lipocalines humaines dans l'ensemble des tissus étudiés (Figure 10). Ceci a conduit à montrer que le gène codant pour l'apolipoprotéine D est exprimé dans les glandes du cornet nasal et du méat moyen, en faisant ainsi une 20 potentielle odorant-binding protéine. De même la rétinol-binding protéine (RBP) est exprimée par l'épithélium nasal et pourrait aussi participer à cette fonction.

Les inventeurs proposent donc d'inclure les protéines ApoD et RBP dans la famille des OBP humaines et les incluent dans 25 l'ensemble des revendications portant sur ces OBP humaines.

Les inventeurs ont démontré que la transcription des deux gènes hOBPII génère de nombreux transcrits alternatifs qui codent pour des protéines distinctes dont la structure est compatible avec celle d'un transporteur de ligand hydrophobe.

Les inventeurs ont mis en évidence l'expression des deux gènes hOBPII dans la sphère orale (glandes nasales, glandes de von Ebner, glandes sous-maxillaires, glandes lacrymales, poumon). Les inventeurs ont également démontré que les protéines hOBPII-selon
5 l'invention sont produites par les cellules de la sphère génitale ; le gène hOBPII est majoritairement exprimé dans la prostate, le canal déférent et les glandes mammaires alors que l'expression du gène hOBPIIa est restreinte au canal déférent. Les analyses in situ ont
10 révélé que l'ARNm est produit par les cellules glandulaires de la prostate et des cellules sécrétrices épithéliales du canal déférent supportant l'idée d'une sécrétion des protéines correspondantes dans le fluide séminal. Les protéines selon l'invention sont donc susceptibles d'être impliquées dans la fonction de reproduction, mais également dans toutes les autres fonctions habituellement
15 attribuées aux lipocalines.

EXEMPLE 7 : FABRICATION D'UNE PROTEINE RECOMBINANTE HOBPIIb α DANS UN SYSTEME PROCARYOTE.

20

Une PCR sur l'ADN plasmidique d'un clone de hOBPIIb α à l'aide des amorces BIIa/b (5' GTC GGA TCC CTG TCC TTC ACC CTG GAG G 3'), oligonucléotide sens démarrant 45 bases après l'ATG initiant la protéine, et XIIb (5' GTC CTC GAG GTG TTC GGG
25 AAC GCA GCT TC 3'), oligonucléotide antisens précédant le codon stop de la protéine hOBPIIb α , a permis d'amplifier la totalité de l'ADN codant pour la protéine hOBPIIb α sécrétée. Les sites de restriction enzymatique BamH I et Xho I situés aux extrémités des deux oligonucléotides (bases soulignées), furent utilisés pour un

clonage directionnel dans un vecteur plasmidique d'expression pGEX-6P1, suivi d'une transformation par électroporation (1800V, 200 Ω , 25 μ F) dans une souche bactérienne BL21. La synthèse de la protéine recombinante est obtenue par ajout d'IPTG à raison de 5 mM final pendant 3 h dans 250 ml de culture de la souche en milieu LB contenant 100 μ g/ml d'ampicilline préalablement incubée à 37°C pendant 2 h. Les cultures centrifugées sont reprises dans 25 ml de tampon TENG. Le lysat est soniqué et centrifugé à nouveau. La protéine de fusion est alors purifiée à l'aide de 4 ml de billes fixant de façon covalente du glutathion insoluble (Sigma) pour 25 ml de surnageant. Après 4 h d'incubation, elles sont lavées avec 3 volumes de NaCl 1M puis par 10 volumes de PBS 1X. L'élution est obtenue par mise en contact des billes pendant 10 min avec une solution de glutathion (Glutathion réduit 10 mM, Tris-HCl pH 8,0 50 mM).

La quantité de protéine recombinante produite est estimée par migration en gel de polyacrylamide et coloration au bleu de Coomassie (Figure 11). La spécificité de l'induction d'une protéine recombinante est testée par Western blot à l'aide d'un anticorps anti-GST de chèvre (Figure 11), révélé avec un anticorps anti-IgG de chèvre couplé à la peroxydase en utilisant un kit ECL+plus (Amersham). La séparation de la protéine de fusion en deux protéines est obtenue par protéolyse de 100 μ g de protéine recombinante préalablement dialysée à l'aide de 2U de preScissionTM protéase (Pharmacia Biotech) dans 10 μ l final contenant du tampon de l'enzyme 1X, pendant 4 h à 5°C.

**EXEMPLE 8 : DETECTION DES hOBPII DANS DIFFERENTS
TISSUS PAR IMMUNOHISTOCHIMIE.**

La localisation des protéines hOBP_{II} au sein des structures nasales (figures 12 et 15), des structures buccales et des glandes lachrymales (figures 13 et 18), des glandes mammaires et des poumons (figures 14 et 18) est révélée par immunohistochimie.

5

EXEMPLE 9 : DETECTION DES hOBP_{II}, LCN1 ET ApoD, AU SEIN DES STRUCTURES NASALES PAR RT-PCR, ET PAR IMMUNOHISTOCHIMIE.

10 Les RT-PCR réalisées à l'aide des jeux d'amorces pour les hOBP_{II} et pour LCN1 révèlent la présence des ARNm correspondants dans au moins une des structures nasales. Les techniques de détection immunohistochimiques tant par détection de fluorescence que par révélation à l'aide de peroxydase et de DAB, 15 révèlent la présence des protéines hOBP_{II} (figure 15) et LCN1 (figure 16) dans les cellules épithéliales des acini des glandes nasales, ainsi que dans les lumières des canaux sécréteurs. On peut également en détecter sur du mucus collé sur les tissus.

Les RT-PCR réalisées à l'aide des jeux d'amorces pour 20 l'ApolipoprotéineD révèlent la présence des ARNm correspondants dans au moins une des structures nasales. Les techniques de détection immunohistochimiques par révélation à l'aide de peroxydase et de DAB, révèlent la présence de la protéine ApoD (figure 17) dans les cellules épithéliales des acini des glandes 25 nasales, ainsi que dans les lumières des canaux sécréteurs, bien que l'expression semble plus faible et dans moins d'acini que pour les hOBP_{II}.

**EXEMPLE 10 : DETECTION DES hOBPII, LCN1, ApoD DANS LE
PLACENTA PAR RT-PCR ET PAR
IMMUNOHISTOCHIMIE.**

Les RT-PCR réalisées à l'aide des jeux d'amorces pour les
5 hOBPII et pour LCN1 révèlent la présence des ARNm
correspondants dans le placenta. Les techniques de détection
immunohistochimiques par révélation à l'aide de peroxydase et de
DAB, révèlent la présence des protéines hOBPII (figure 19) et LCN1
10 (figure 20) très fortement au niveau des lacunes des villosités
placentaires et beaucoup plus faiblement et de façon très diffuse au
niveau des cellules épithéliales de la membrane amniotique et de
cellules sous jacentes tant pour la membrane amniotique que pour
le cordon. L'expression des ARNm détectés par RT-PCR au niveau
des villosités placentaires indique que ces protéines hOBPII sont
15 produites par le placenta pour transférer des molécules lipophiles
(hormones, acides gras, vitamines...) vers le fœtus par le biais de la
circulation sanguine fœtale. Les protéines détectées au niveau de la
membrane amniotique pourraient correspondre à l'absorption de
ces protéines à partir du liquide amniotique, sachant que les
20 structures nasales du fœtus sont baignées par ce liquide
amniotique. Les protéines hOBPII et LCN1 participeraient ainsi au
transport de substances lipophiliques de petite taille de la mère
vers le fœtus au niveau des villosités choriales par l'intermédiaire
du sang foetal, et du fœtus vers la mère par l'intermédiaire du
25 liquide amniotique.

Les RT-PCR réalisées à l'aide des jeux d'amorces pour la
protéine ApoD révèlent la présence des ARNm correspondants dans
le placenta. Les techniques de détection immunohistochimiques par
révélation à l'aide de peroxydase et de DAB, révèlent la présence de

la protéine ApoD très fortement au niveau des lacunes des villosités placentaires (figure 21). La protéine ApoD n'est pas détectée au niveau des cellules de la membrane amniotique et de cellules sous jacentes tant pour la membrane amniotique que pour le cordon.

- 5 L'expression des ARNm détectés par RT-PCR au niveau placentaire indique que la protéine ApoD est produite par le placenta pour transférer des molécules lipophiles (hormones, acides gras, vitamines...) vers le fœtus par le biais de la circulation sanguine fœtale. La protéine ApoD peut également être produite par le foie du
- 10 fœtus. La protéine ApoD participerait ainsi au transport de substances lipophiliques de petite taille de la mère vers le fœtus au niveau des villosités choriales par l'intermédiaire du sang foetal.

- Plus généralement, l'expression de lipocalines par le placenta, telles que l'Alpha1-acid Glycoprotéine (A1AG), la
- 15 prostaglandine D-synthase (PGDS), PP14 et RBP comme révélée par les techniques de RT-PCR indique que ces protéines aux propriétés de transporteur de molécules hydrophobes sont massivement recrutées pour les relations mère-enfant, pour les transports de ces molécules comme peuvent être des hormones stéroïdes, des acides
- 20 gras, des vitamines ou des xénobiotiques et plus généralement des produits de détoxification issus du métabolisme.

REFERENCES

- Attwood, J. *et al.* (1994) *Genomics* 19, 203-214.
- 5 Barany, F., (1991), *Proc. Natl. Acad. Sci. USA*, 88, 189-193.
- Beste *et al.* (1999), *Proc. Natl. Acad. Sci. USA*, 96, 1898-1903.
- Bocskai, Z. *et al.* (1992) *Nature* 360, 186-188.
- 10 Borghoff S.J. *et al.* (1990) *Annu. Rev. pharmacol. Toxicol.* 30, 349-367
- Burg, J.L. *et al.* (1996), *Mol. and Cell. Probes*, 10, 257-271.
- 15 Chu, B.C.F. *et al.* (1986), *Nucleic Acids Res.*, 14, 5591-5603.
- Dear, T. N. *et al.* (1991) *Biochemistry* 30, 10376-10382.
- 20 Dewald, G. *et al.* (1996) *Ann Hum Genet* 60, 281-291.
- Duck, P. *et al.* (1990), *Biotechniques*, 9, 142-147.
- Edwards, C.P., and Aruffo, A., (1993), *Curr. Op. Biotechnology* 4,
25 558-563.
- Erlich, H.A., (1989), New York : Stockton Press.
- Flower, D. R. (1995) *J Mol Recognit* 8, 185-195.
- 30 Flower, D. R. (1996) *Biochem J* 318, 1-14.

- Gilbertson T.A. (1998) *Cur. Op. Neurobiology* 8, 447-452.
- Gregoire, C. *et al.* (1996) *J Biol Chem* 271, 32951-32959.
- 5 Guatelli J.C. *et al.* (1990), *Proc. Natl. Acad. Sci. USA*, 87, 1874-1878.
- Henzel, W. J. *et al.* (1988) *J Biol Chem* 263, 16682-16687.
- 10 Holzfeind, P. and Redl, B. (1994) *Gene* 139, 177-183.
- Hornigold, N. *et al.* (1997) *Genomics* 41, 385-389.
- Innis, M.A. *et al.* (1990), Academic Press.
- 15 Igarashi, M. *et al.* (1992) *Proc Natl Acad Sci U S A* 89, 5376-2380.
- Kievitis, T. *et al.* (1991), *J. Virol. Methods*, 35, 273-286.
- 20 Kock, K. *et al.* (1994) *Eur J Biochem* 221, 905-916.
- Kohler, G. *et al.* (1975), *Nature*, 256 (5517), 495-497.
- Konieczny, A. *et al.* (1997) *Immunology* 92, 577-586.
- 25 Kwoh, D.Y. *et al.* (1989), *Proc. Natl. Acad. Sci. USA*, 86, 1173-1177.
- Lacazette, E. *et al.* (1997) *Ann Hum Genet* 61, 449-455.
- 30 Landegren, U. *et al.* (1988), *Science*, 241, 1077-1080.

- Lassagne, H. and Gachon, A. M. (1993) *Exp Eye Res* 56, 605-609.
- Lassagne, H. *et al.* (1993) *Genomics* 18, 160-161.
- 5 Lassagne, H. *et al.* (1995) *Cytogenet Cell Genet* 71, 104.
- Lizardi, P.M. *et al.* (1988) *Biotechnology* 6, 1197-1202.
- Luckow, V.A. (1993) *Curr. Op. Biotechnology* 4, 564-572.
- 10 Matthews, J.A. *et al.* (1988) *Anal. Biochem.* 169 : 1-25.
- Miele, E.A. *et al.* (1983) *J. Mol. Biol.* 171 : 281-295.
- 15 Miller, W. L. (1998) *Clin Perinatol* 25, 799-817, v.
- Monaco, H. L. and Zanotti, G. (1992) *Biopolymers* 32, 457-465.
- Nagata, A. *et al.* (1991) *Proc Natl Acad Sci U S A* 88, 4020-4024.
- 20 Nahmias, J. *et al.* (1995) *Eur J Hum Genet* 3, 65-77.
- Pelosi, P. (1996) *J Neurobiol* 30, 3-19.
- 25 Pervaiz, S. and Brew, K. (1987) *Faseb J* 1, 209-214.
- Redl, B. *et al.* (1992) *J Biol Chem* 267, 20282-20287.
- Rolfs, A., *et al.* (1991), Berlin : Springer-Verlag.

Sambrook, J. *et al.* (1989), T. Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

Segev, D., (1992), Kessler C. Springer Verlag, Berlin, New-York,
5 197-205.

Senoo, H. *et al.* (1990) J Lipid Res 31, 1229-1239.

Shaw, P. H. *et al.* (1983) Cell 32, 755-761.
10

Simard *et al.* (1992) Endocrinology 30, 1115-1121.

Spinelli, S. *et al.* (1998) Biochemistry 37, 7913-7918.

15 Stoesz S.P. and Gould M.N. (1995) Oncogene 11: 2233-2241.

Stone, B.B. *et al.* (1996) Mol. and Cell. Probes 10 : 359-370.

van Slegtenhorst, M. *et al.* (1995) Eur J Hum Genet 3, 78-86.
20

Walker, G.T. *et al.* (1992) Nucleic Acids Res. 20 : 1691-1696.

Zeng, C. *et al.* (1996) Proc Natl Acad Sci U S A 93, 6626-6630.

REVENDICATIONS

1. Polypeptide isolé comprenant une séquence en amino-
acides ayant au moins 90% d'identité avec les séquences en amino-
5 acides SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ
ID N° 12 ou SEQ ID N° 14.

2. Polypeptide isolé caractérisé en ce qu'il comprend un
polypeptide choisi parmi :

10 a) un polypeptide de séquence SEQ ID N° 2, SEQ ID N° 4,
SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 12 ou SEQ ID N° 14;

b) un polypeptide variant de polypeptide de séquences
d'acides aminés défini en a) ;

c) un polypeptide homologue au polypeptide défini en a) ou
15 b) et comportant au moins 90 % d'identité avec ledit polypeptide de
a) ;

d) un fragment d'au moins 15 acides aminés consécutifs
d'un polypeptide défini en a), b) ou c) ;

e) un fragment biologiquement actif d'un polypeptide défini
20 en a), b) ou c).

3. Polypeptide isolé sélectionné parmi un polypeptide
correspondant à la séquence SEQ ID N°2 et dénommé OBPII_{aa} , à la
séquence SEQ ID N°4 et dénommé OBPII_{ap} , à la séquence SEQ ID
25 N°6 et dénommé OBPII_{ay}, à la séquence SEQ ID N° 12 et dénommé
OBPII_{bp} .

4. Polypeptide selon l'une quelconque des revendications 1 à 3 caractérisé en ce qu'il comporte au moins le domaine Gly-Thr-Trp-Tyr.

5 5. Polynucléotide isolé caractérisé en ce qu'il code pour un polypeptide selon l'une quelconque des revendications 1 à 4.

6. Polynucléotide selon la revendication 5 de séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 11 ou
10 SEQ ID N° 13.

7. Polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

a) un polynucléotide de séquence SEQ ID N° 1, SEQ ID N°
15 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 11 ou SEQ ID N° 13 ou dont la séquence est celle de l'ARN correspondant à la séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 11 ou SEQ ID N° 13 ;

b) un polynucléotide dont la séquence est complémentaire
20 de la séquence d'un polynucléotide défini en a),

c) un polynucléotide dont la séquence comporte au moins 80% d'identité, avec un polynucléotide défini en a) ou b),

d) un fragment d'au moins 15 nucléotides consécutifs, de préférence 21 nucléotides consécutifs, et de manière préférée 30
25 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).

8. Polynucléotide selon la revendication 7 caractérisé en ce qu'il est marqué directement ou indirectement par un composé radioactif ou un composé non radioactif.

9. Utilisation d'un polynucléotide selon la revendication 8 en tant que sonde pour la détection de séquences nucléiques.

10. Utilisation d'un polynucléotide selon l'une quelconque
5 des revendications 5 à 8 en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques.

11. Utilisation d'un polynucléotide selon la revendication 8
10 en tant que oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant.

12. Vecteur recombinant de clonage d'un polynucléotide selon l'une des revendications 5 à 8 et/ou d'expression d'un polypeptide selon l'une des revendications 1 à 4 caractérisé en ce
15 qu'il contient un polynucléotide selon l'une quelconque des revendications 5 à 8.

13. Vecteur selon la revendication 12 caractérisé en ce qu'il comporte les éléments permettant l'expression éventuellement la
20 sécrétion dudit polypeptide dans une cellule hôte.

14. Vecteur selon l'une quelconque des revendications 12 à 13 caractérisé en ce que les éléments permettant l'expression dudit polypeptide sont choisis parmi :

25 a) le polynucléotide isolé de séquence SED ID N°15 et SEQ ID N° 16;

b) un polynucléotide dont la séquence est complémentaire de la séquence du polynucléotide défini en a) ;

c) un polynucléotide dont la séquence comporte au moins 80% d'identité avec un polynucléotide défini en a) ou en b) ;

d) un polynucléotide hybridant dans des conditions de forte stringence avec une séquence du polynucléotide défini en a), b) ou
5 c).

15. Vecteur selon les revendications 13 et 14 pour l'expression dans les cellules eucaryotes sélectionnés parmi l'ADN viral et l'ADN nu.

10

16. Cellule hôte, caractérisée en ce qu'elle est transformée par un vecteur selon l'une des revendications 12 à 15.

17. Procédé de préparation d'un polypeptide recombinant
15 caractérisé en ce que l'on cultive une cellule hôte selon la revendication 16 dans des conditions permettant l'expression et éventuellement la sécrétion dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.

20 18. Polypeptide recombinant susceptible d'être obtenu par un procédé selon la revendication 17.

19. Utilisation d'un polypeptide choisi parmi un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 ou l'un de ses
25 fragments, comme protéine de liaison à un ligand hydrophobe de préférence une molécule odorante.

20. Utilisation d'un polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou du polypeptide hOBPII_{bx} de séquence

SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba},
un polypeptide homologue audit polypeptide hOBPII_{ba} comportant
au moins 90% d'identité avec ledit polypeptide hOBPII_{ba} ou l'un de
ses fragments, comme inhibiteur compétitif comme, agoniste ou
5 antagoniste des récepteurs cellulaires aux lipocalines.

21. Anticorps monoclonal ou polyclonal et ses fragments
caractérisé en ce qu'il est dirigé spécifiquement contre un
polypeptide isolé selon l'une des revendications 1 à 4 et 18.

10

22. Utilisation d'un anticorps selon la revendication 21 pour
la mise en évidence de la présence d'un polypeptide selon l'une des
revendications 1 à 4 et 18 dans un échantillon biologique.

15 23. Procédé de détection d'anticorps dirigé contre hOBPII
dans le sérum humain de patient allergique et/ou asthmatique en
utilisant un polypeptide hOBPII.

20 24. Procédé pour contrôler la volatilisation d'une odeur
caractérisé en ce qu'il comprend une étape de liaison de ladite
odeur avec un polypeptide selon l'une quelconque des
revendications 1 à 4 et 18, ou avec le polypeptide hOBPII_{ba} de
séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide
hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba}
25 comportant au moins 90% d'identité avec ledit polypeptide
hOBPII_{ba}.

25. Procédé selon la revendication 24 caractérisé en ce que
le polypeptide est lié à un support solide.

26. Procédé selon la revendication 24 caractérisé en ce que le polypeptide est dans une composition liquide.

5 27. Procédé selon la revendication 26 caractérisé en ce que ladite composition est une composition parfumée pour la peau.

28. Procédé de criblage de molécule, de préférence les odeurs ou saveurs qui comprend de passer la molécule sur un
10 substrat qui comprend un polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba}, lié audit
15 substrat, ledit polypeptide liant ladite odeur ou saveur et récupérer ladite odeur ou saveur depuis le polypeptide si nécessaire.

29. Procédé selon la revendication 28 caractérisé en ce que les molécules sont des phéromones humaines.

20

30. Procédé pour solubiliser des molécules lipophiles caractérisé en ce qu'il comprend la liaison de ladite molécule à un polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou au polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un
25 polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba}.

31. Application des polypeptides selon l'une des revendications 1 à 4 et 18, ou du polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, d'un polypeptide variant dudit polypeptide hOBPII_{ba}, d'un polypeptide homologue audit polypeptide hOBPII_{ba} comportant
5 au moins 90% d'identité avec ledit polypeptide hOBPII_{ba} en combinaison avec des acides gras alimentaires, à titre d'additif alimentaire.

32. Application selon la revendication 31 pour la
10 préparation d'un médicament destiné au traitement des hyperlipidémies et de l'obésité.

33. Application selon la revendication 31 pour
complémenter les laits non-maternels.
15

34. Polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90%
20 d'identité avec ledit polypeptide hOBPII_{ba} en tant qu'agent de ciblage de composé pharmaceutique.

35. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est exprimé sous la forme d'une protéine de
25 fusion avec une protéine permettant un adressage cellulaire spécifique.

36. Polypeptide selon la revendication 35 caractérisé en ce que ladite protéine permettant un adressage cellulaire spécifique

est choisie dans le groupe composé des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

5 37. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est associé à une molécule permettant un adressage cellulaire spécifique.

10 38. Polypeptide selon la revendication 37 caractérisé en ce que ladite molécule permettant un adressage cellulaire spécifique est choisie dans le groupe composé des stéroïdes, des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

15 39. Polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba} en tant que transporteur
20 de composé pharmaceutique.

40. Composition pharmaceutique comprenant un composé pharmaceutique lié au moins à un polypeptide selon l'une des revendications 34 à 39 et un véhicule pharmaceutiquement
25 acceptable.

41. Composition pharmaceutique selon la revendication 40 caractérisée en ce que le composé pharmaceutique est choisi dans le groupe des agents anticancéreux.

42. Composition pharmaceutique selon la revendication 41 caractérisée en ce que ledit agent anticancéreux est un isotope radioactif choisi parmi le groupe : l'Iode¹³¹, l'Yttrium⁹⁰, l'Or¹⁹⁹, le
5 Palladium¹⁰⁰, le Cuivre⁶⁷, le Bismuth²¹⁷ et l'Antimoine²¹¹.

43. Composition pharmaceutique selon l'une quelconque des revendications 40 à 42 caractérisée en ce que ledit polypeptide selon l'une quelconque des revendications 34 à 39 constitue une
10 forme retard de délivrance dudit composé pharmaceutique dans l'organisme.

44. Composition pharmaceutique comprenant un vecteur d'expression selon les revendications 12 ou 13 et un véhicule
15 pharmaceutiquement acceptable.

45. Composition pharmaceutique selon l'une quelconque des revendications 40 à 44 pour le traitement du cancer choisi de préférence parmi le cancer du sein, le cancer de l'utérus, le cancer
20 de la prostate, le cancer du foie, du carcinome des cellules épithéliales pulmonaires.

46. Utilisation d'un polypeptide choisi parmi le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit
25 polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba}, pour la préparation d'un médicament destiné au traitement du cancer de l'utérus, du cancer de la prostate, du cancer du foie, du cancer des cellules épithéliales pulmonaires.

47. Utilisation du polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, ou un polypeptide variant dudit polypeptide hOBPII_{ba}, ou un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba} pour la préparation d'un médicament destiné au transport de composé à travers la barrière placentaire.

48. Utilisation selon la revendication 47 caractérisée en ce que ledit transport est réalisé de la mère gestante vers le fœtus et que ledit composé est choisi parmi les hormones, les acides gras essentiels, les médicaments lipophiles, les vitamines.

49. Utilisation selon la revendication 47 caractérisée en ce que ledit transport est réalisé du fœtus vers la mère et est destiné à la détoxification du fœtus.

50. Polypeptide selon les revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité avec ledit polypeptide hOBPII_{ba} en tant que marqueur de grossesse.

51. Polypeptide selon les revendications 1 à 4 et 18, ou le polypeptide hOBPII_{ba} de séquence SEQ ID N° 10, un polypeptide variant dudit polypeptide hOBPII_{ba}, un polypeptide homologue audit polypeptide hOBPII_{ba} comportant au moins 90% d'identité

avec ledit polypeptide hOBPII_{ba} en tant que marqueur de pathologie fœto-placentaire.



1 / 21

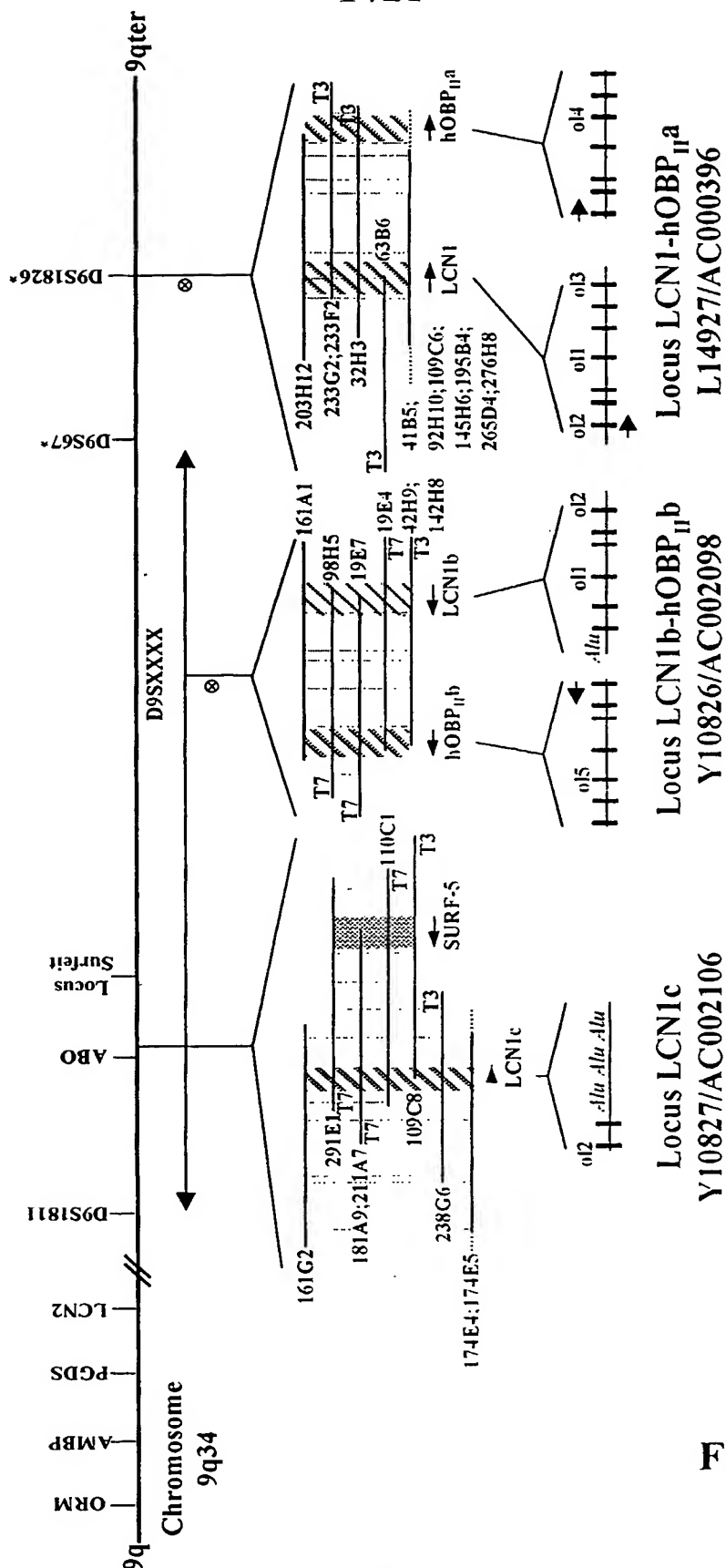


Fig. 1



.

.

.

.

2/21

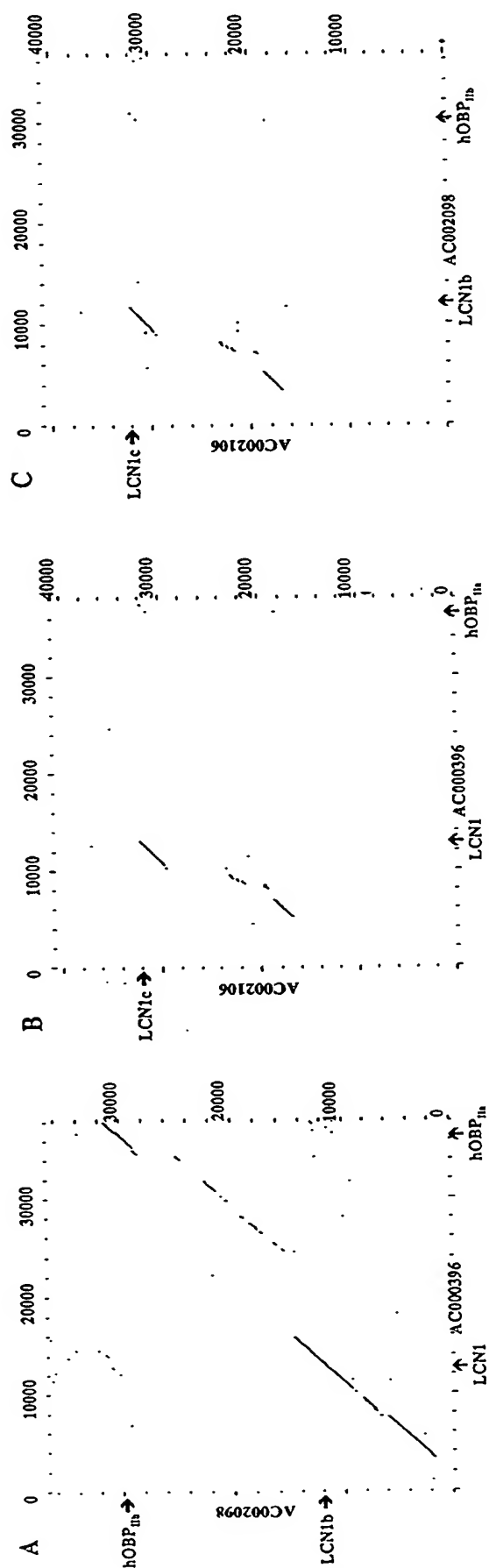


FIG. 2



[illegible]

FIG. 3



.

.

.

5 / 21

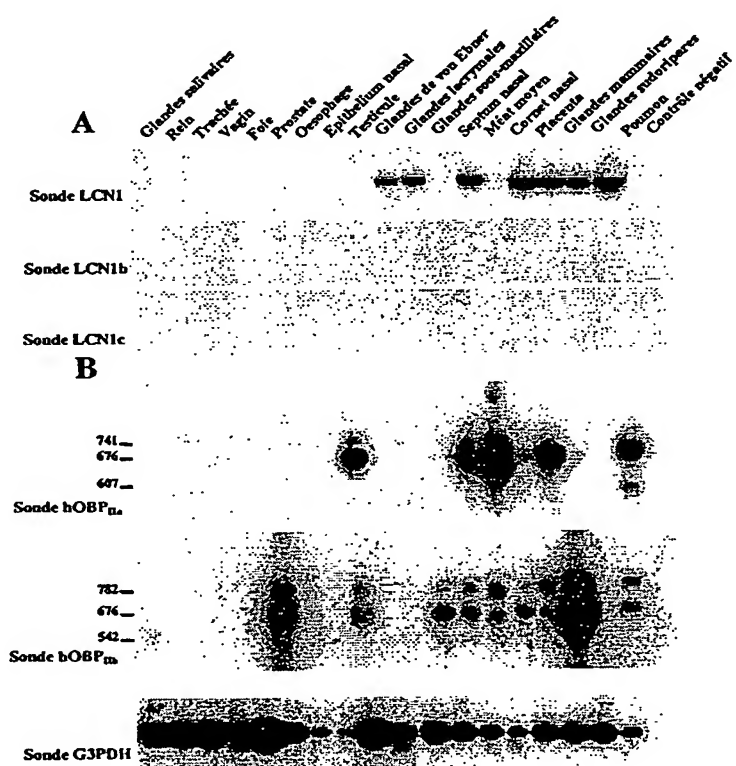
[illegible]

FIG.5



6 / 21

FIG. 6





7 / 21

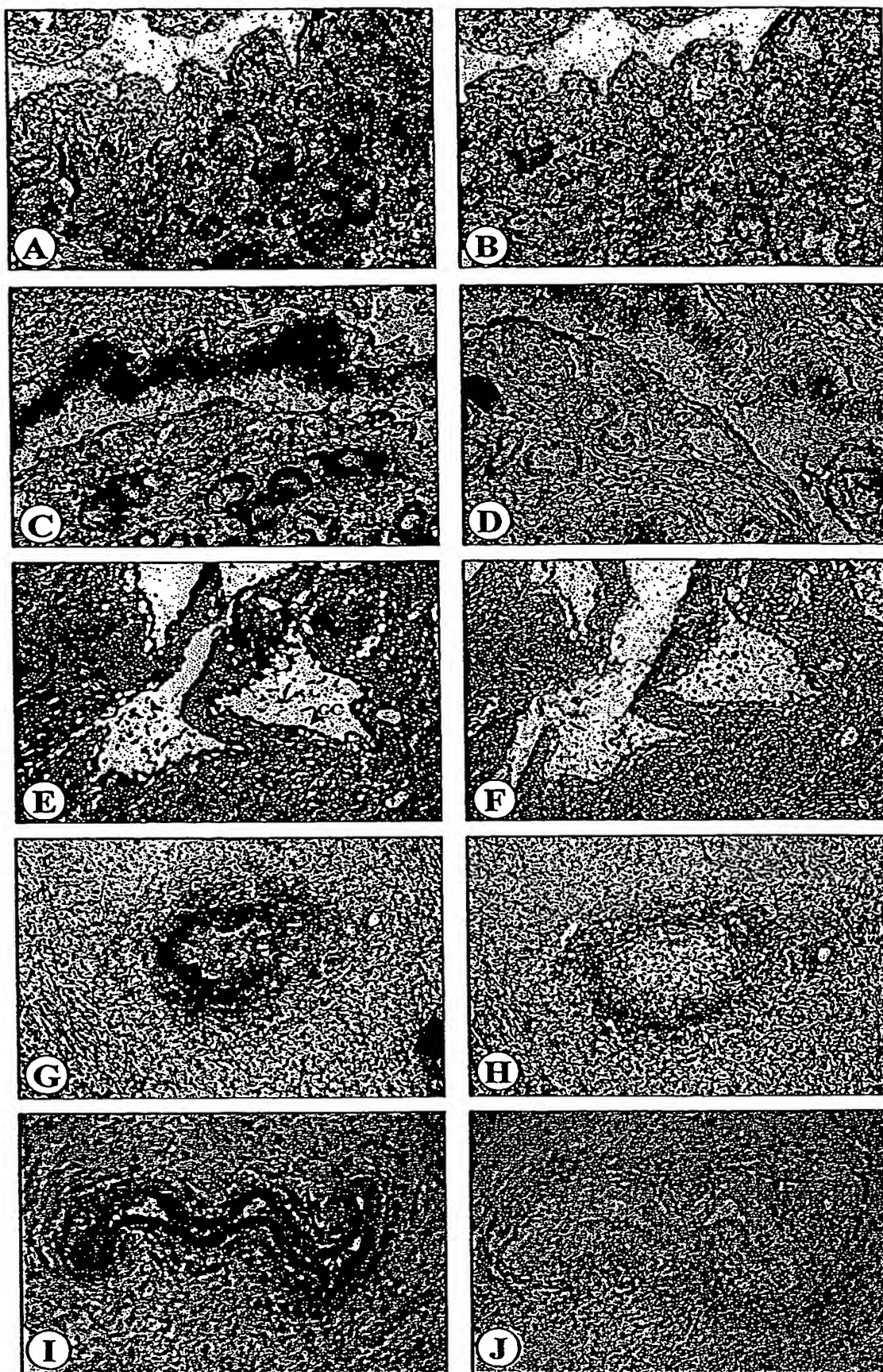


FIG. 7



8 / 21

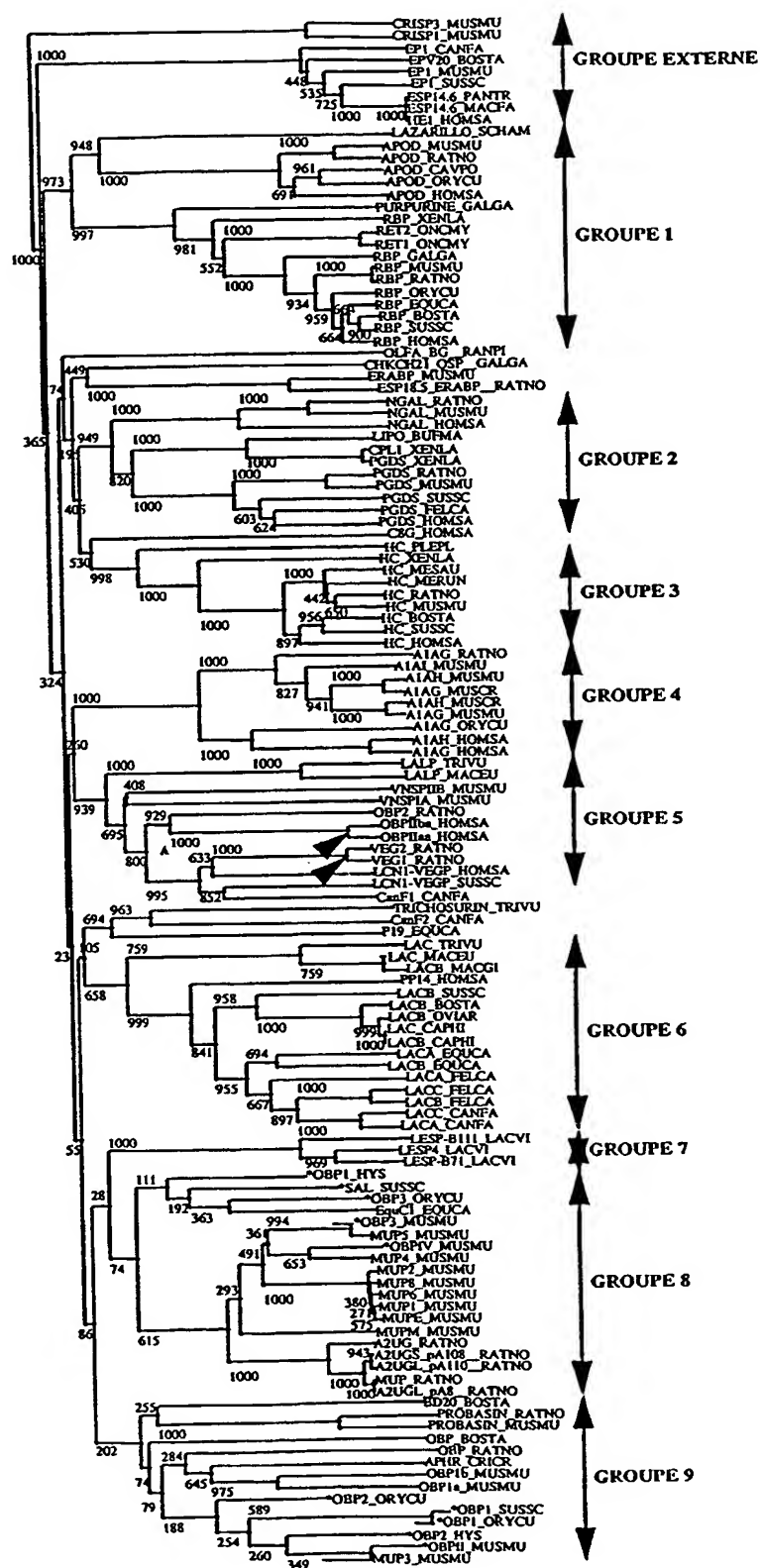


FIG.8



9 / 21

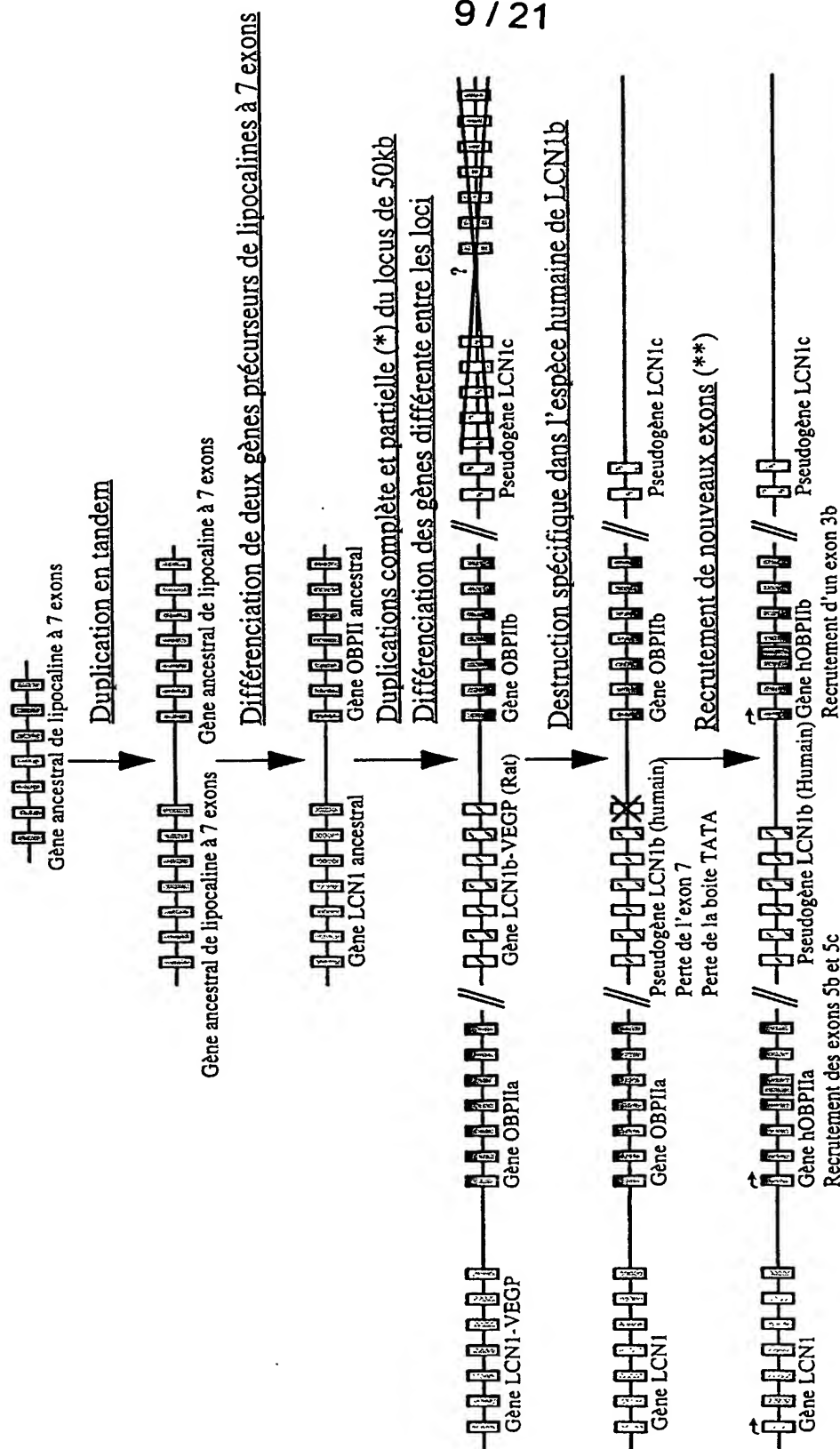


FIG.9



10 / 21

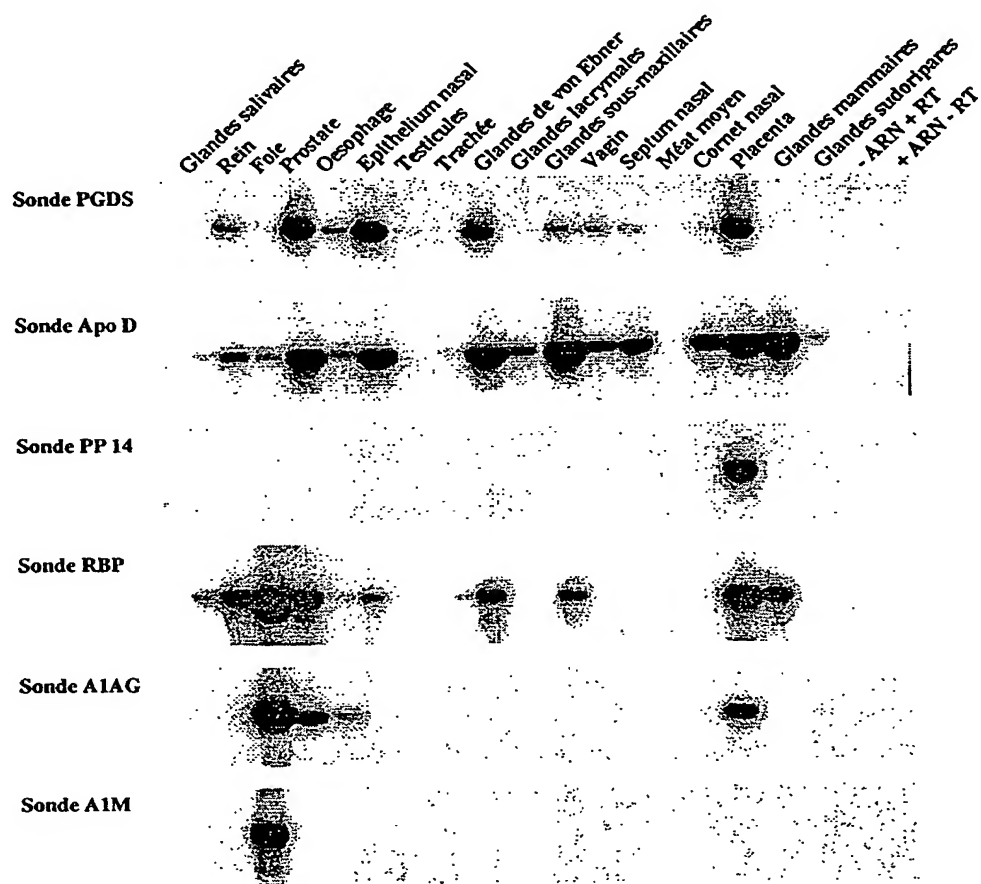


FIG. 10



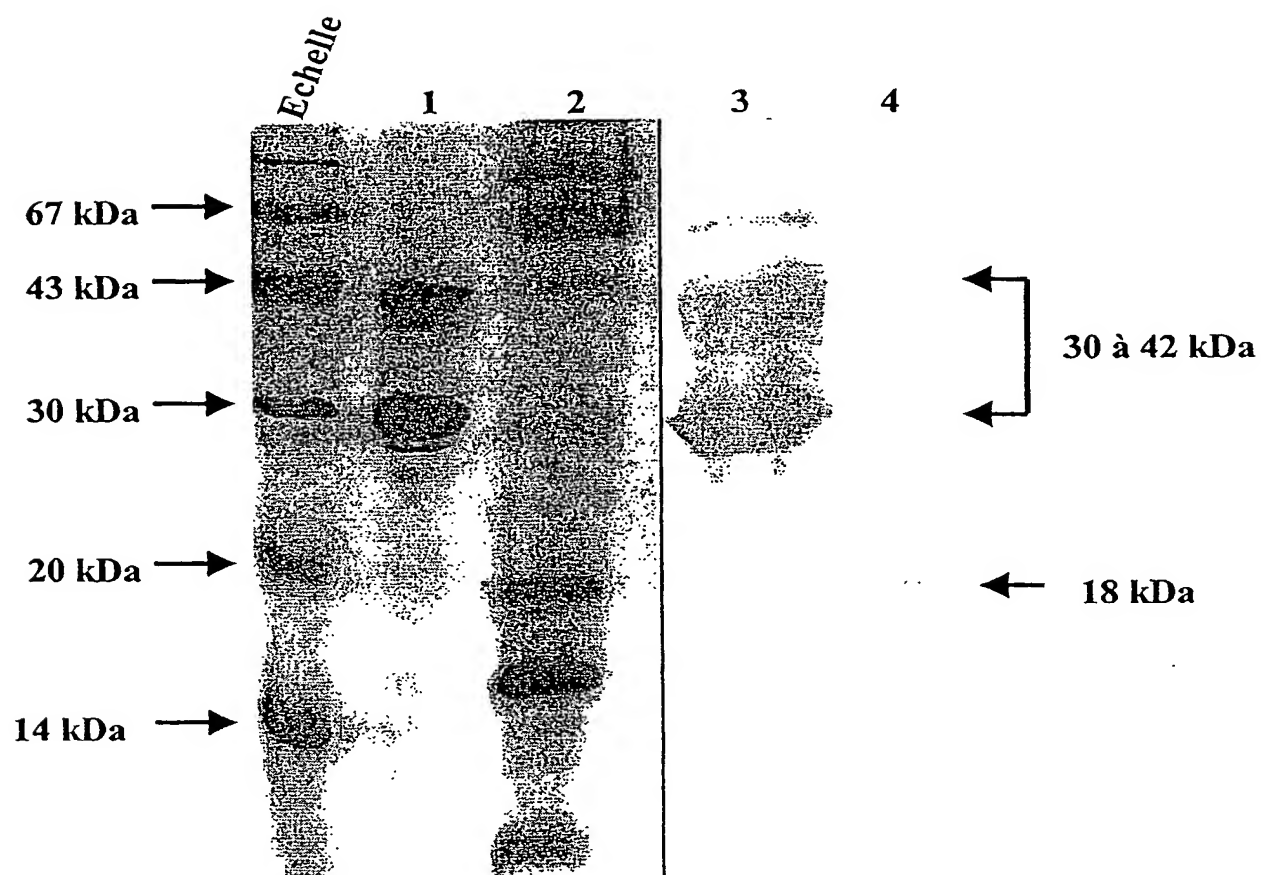


FIG. 11



12 / 21

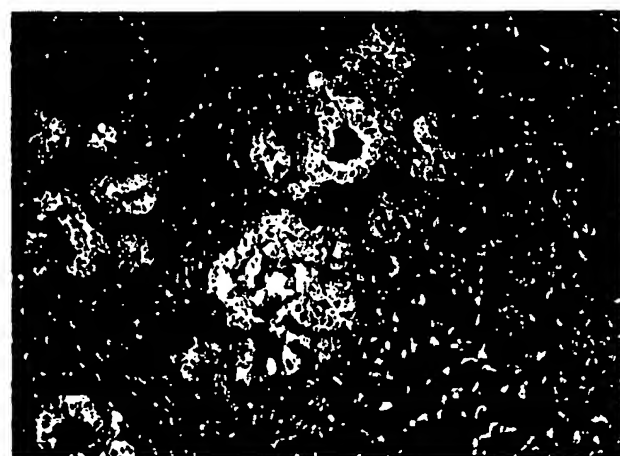
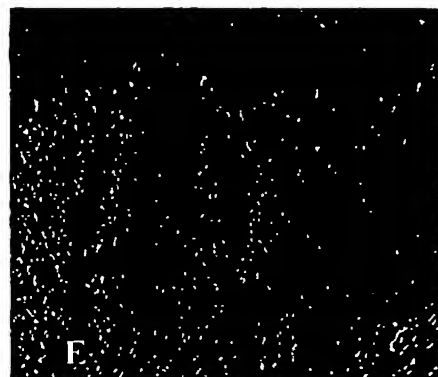
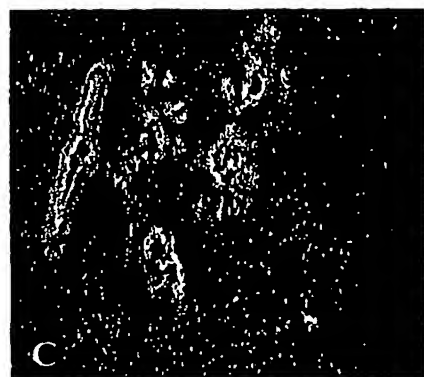
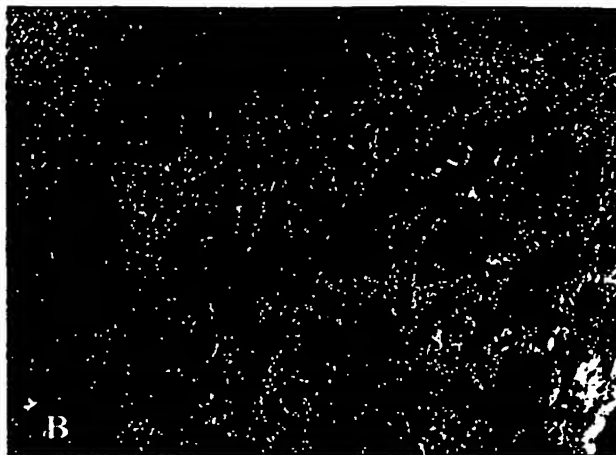


FIGURE 12



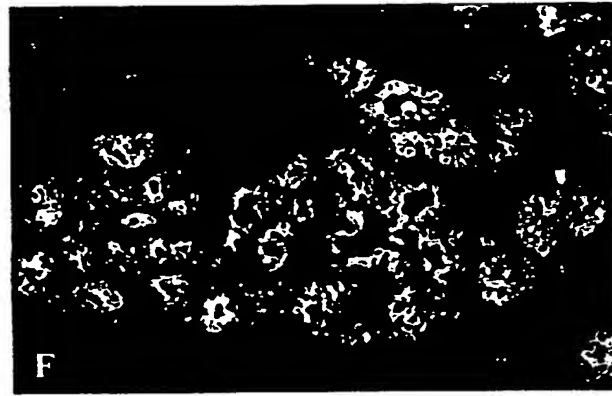
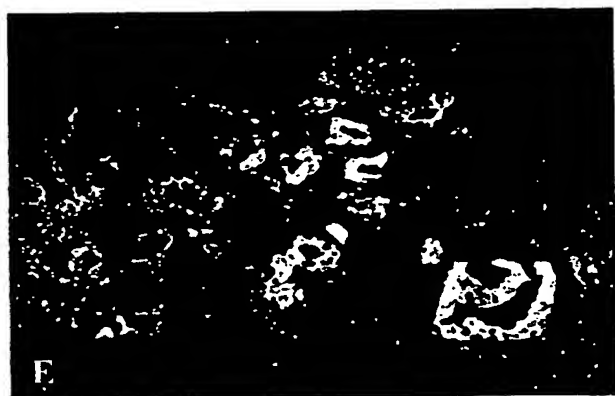
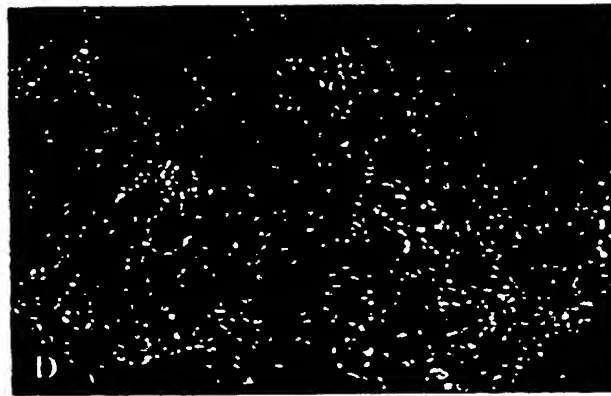
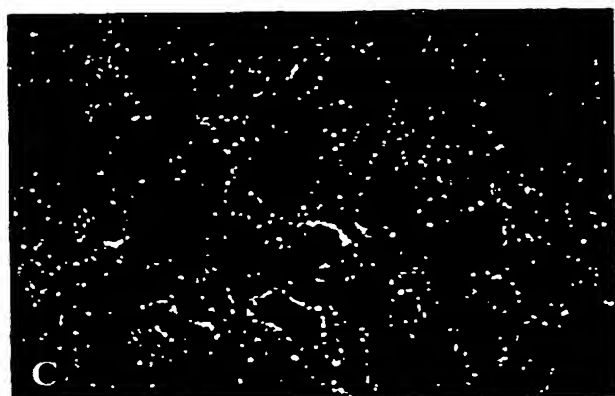
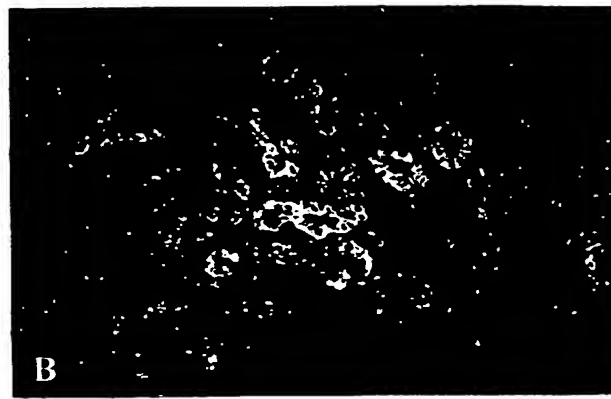
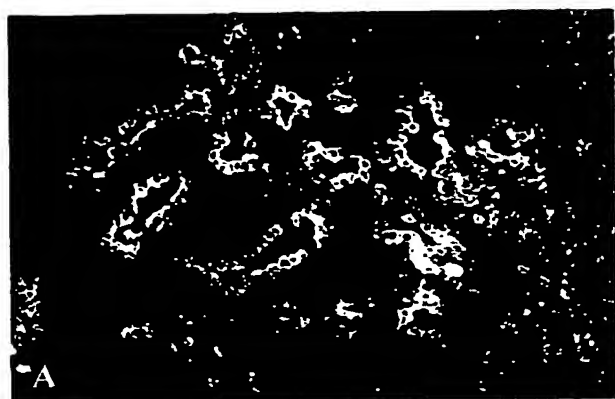


FIGURE 13



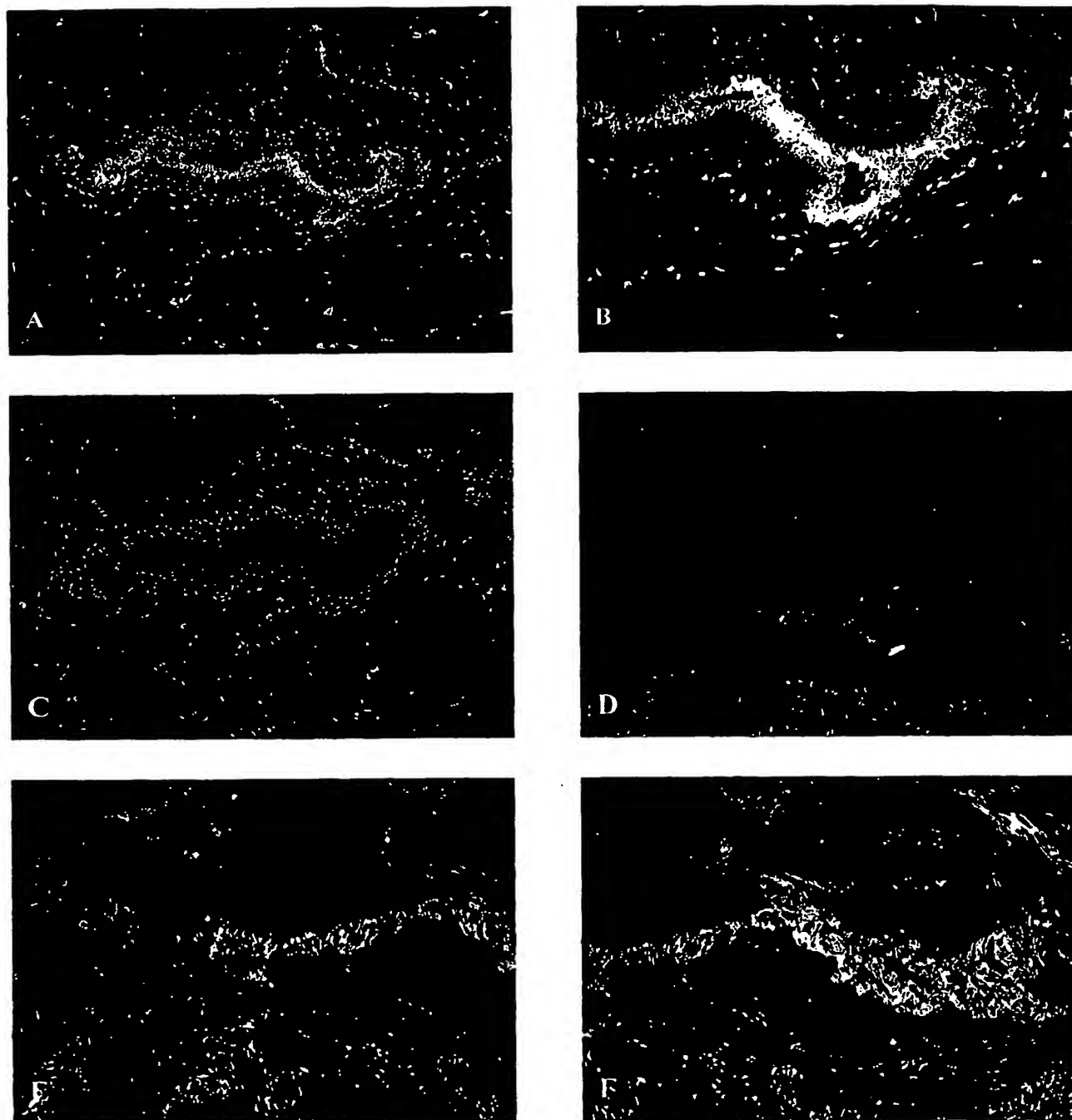


FIGURE 14



15 / 21

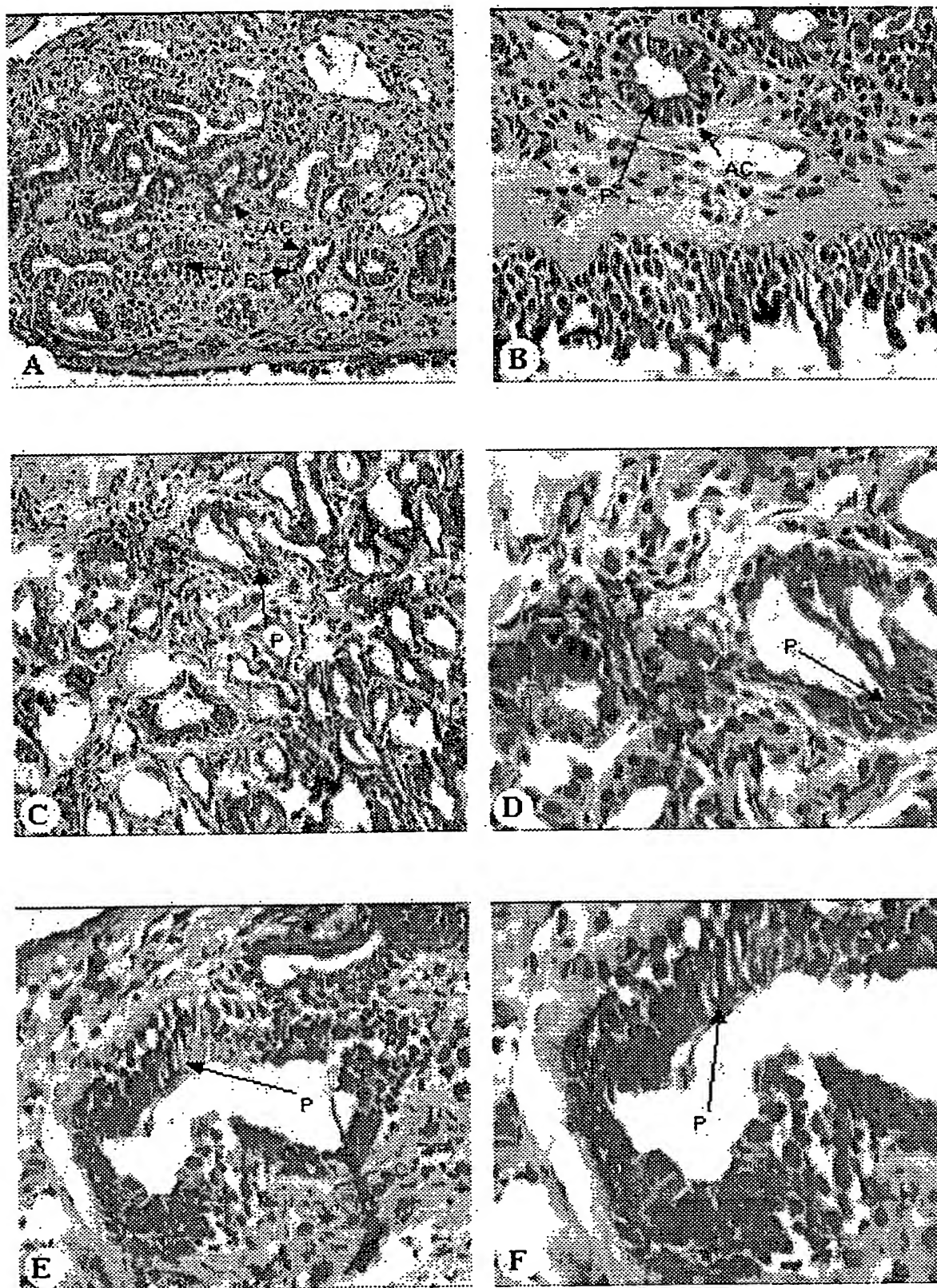


FIG. 15



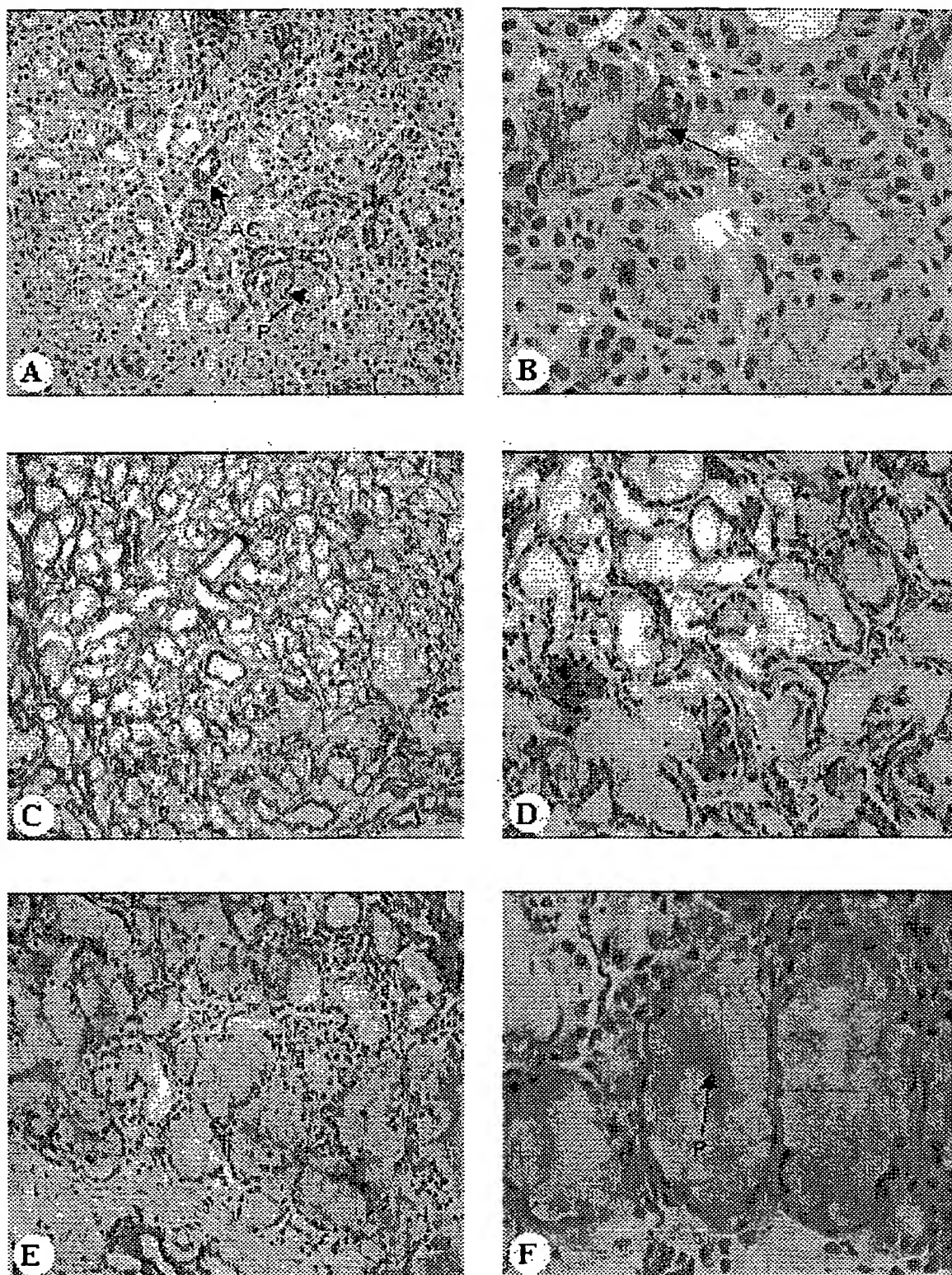


FIG. 16



17 / 21

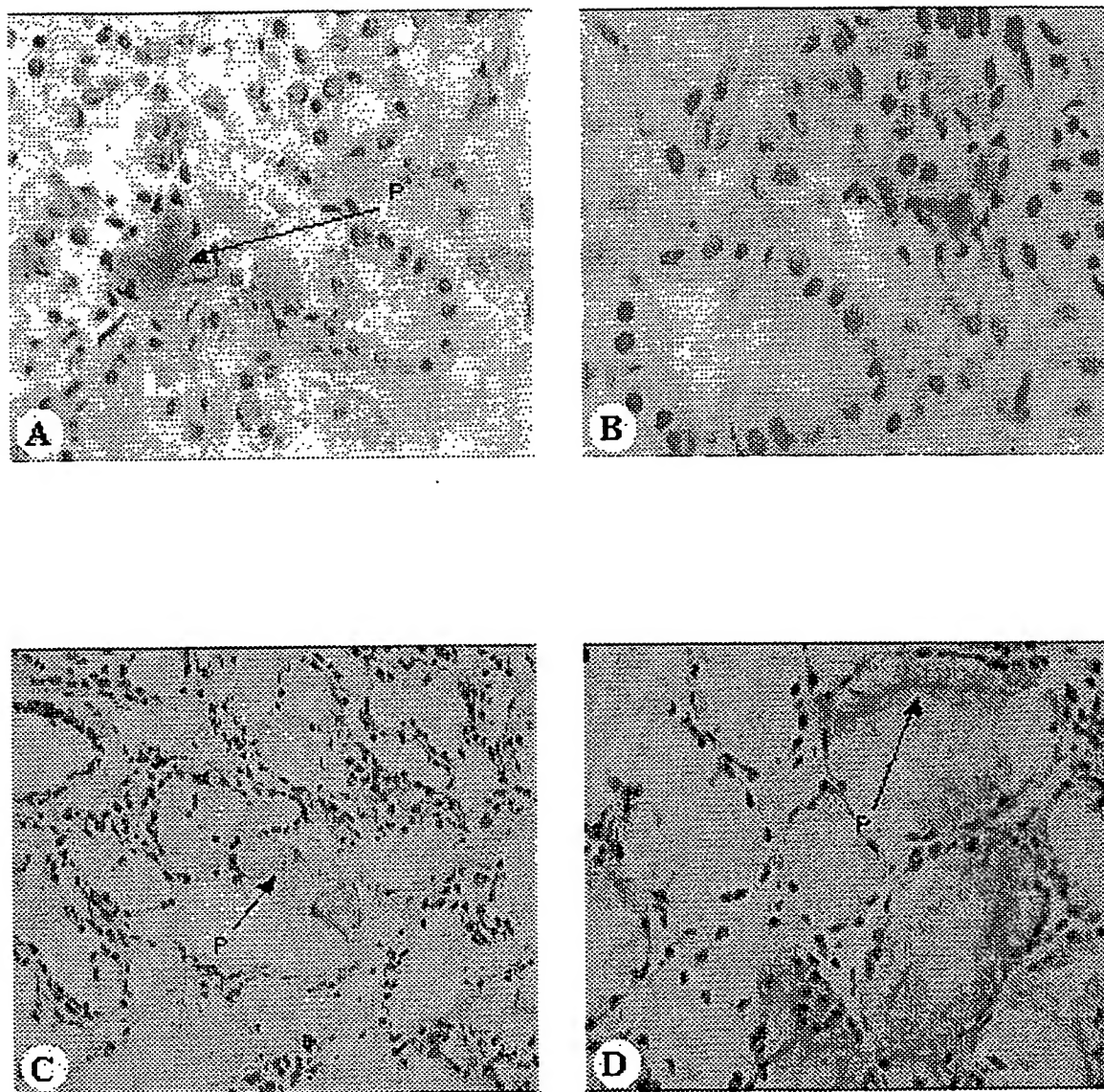


FIG. 17



18 / 21

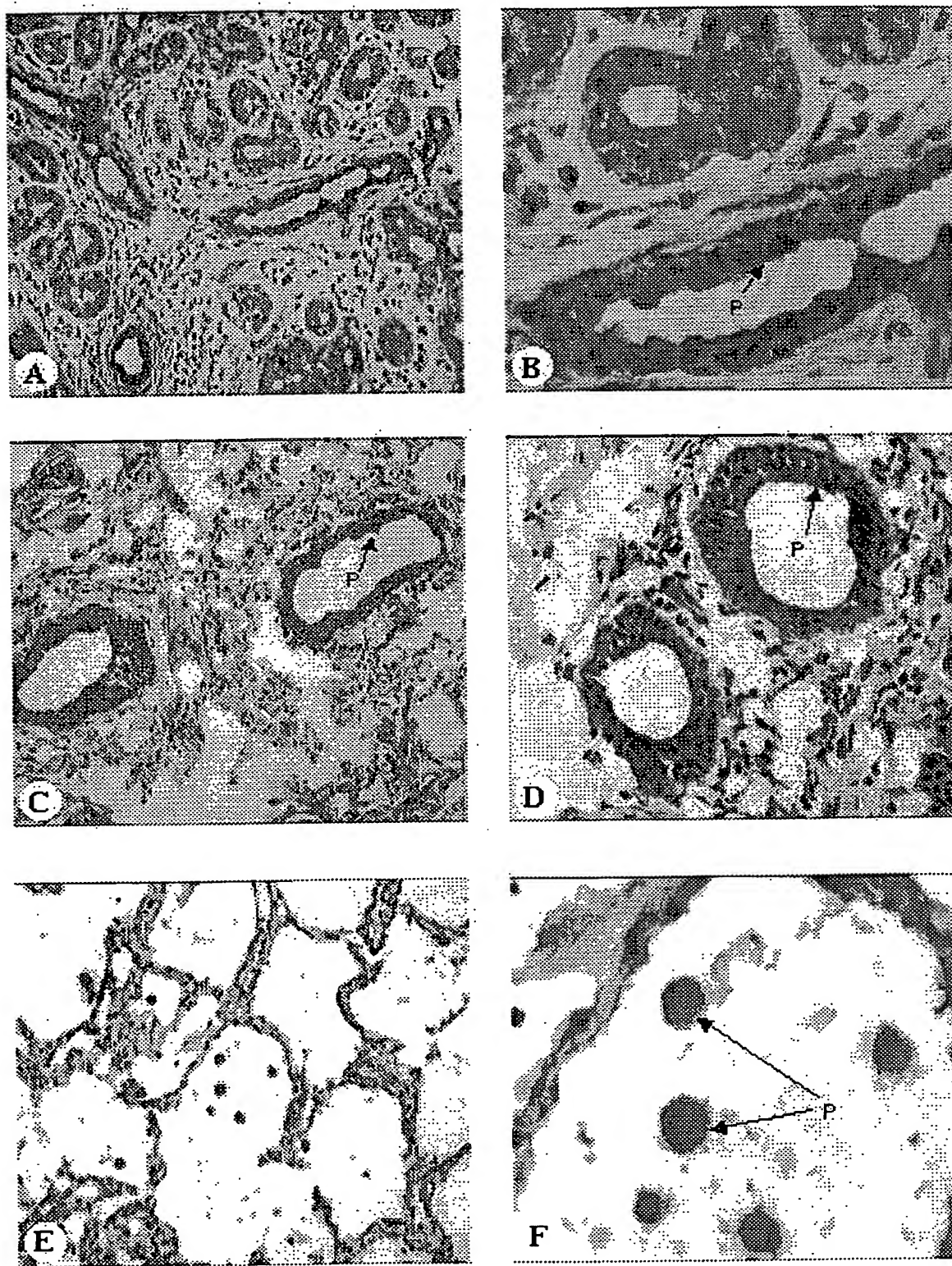


FIG. 18



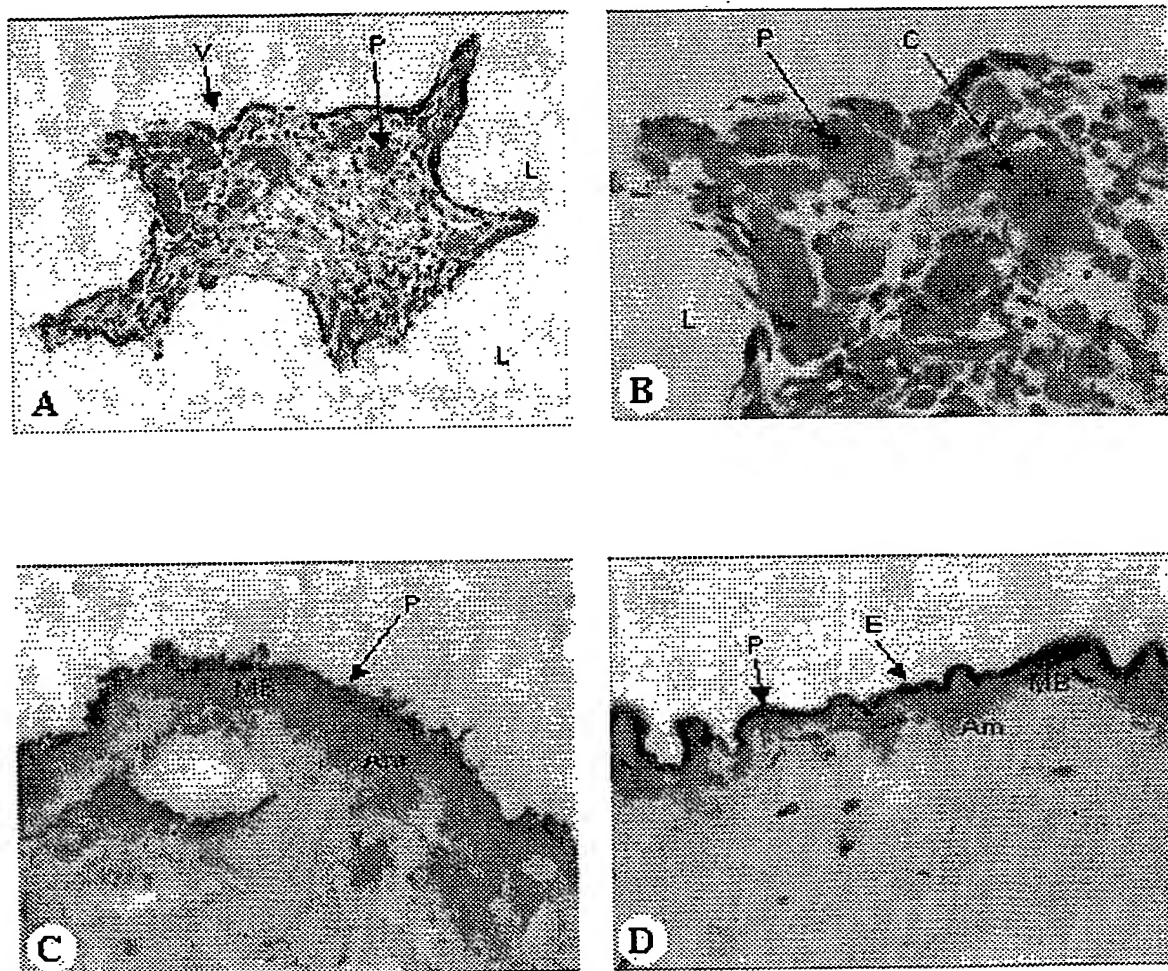


FIG. 19



20 / 21

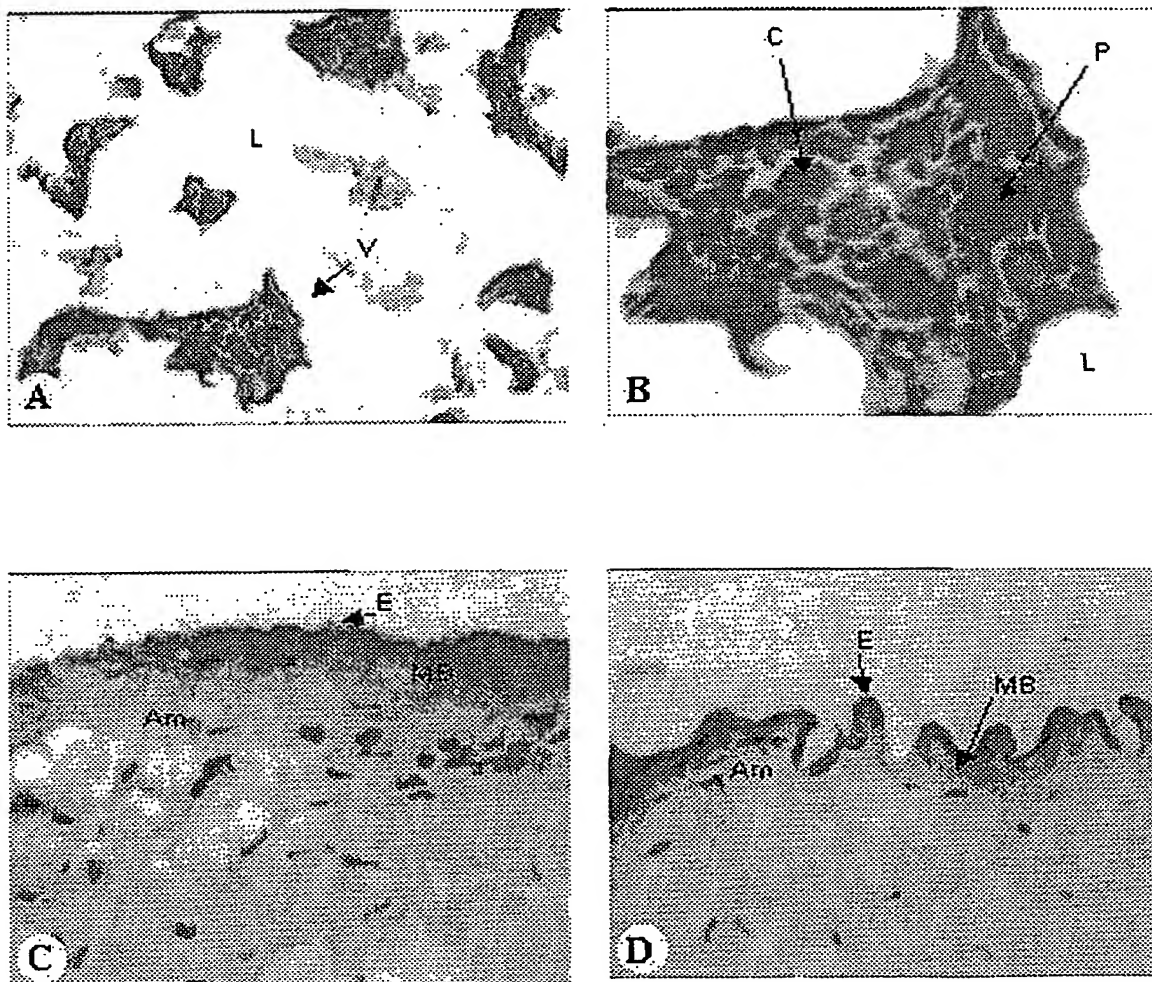


FIG. 20



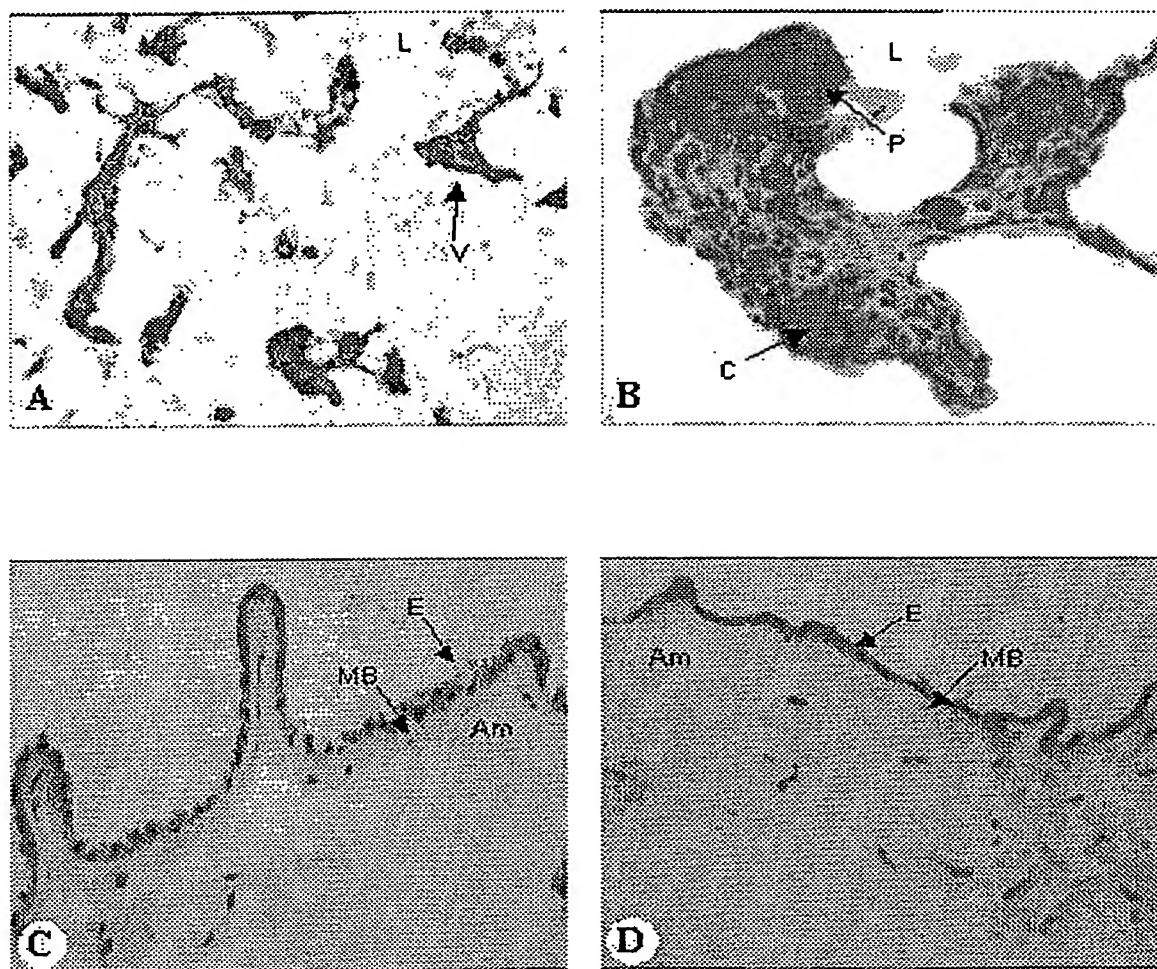


FIG. 21

LISTAGE DE SEQUENCES

<110> UNIVERSITE D'AUVERGNE CLERMONT I

<120> « ODORANT-BINDING » PROTEINES HUMAINES FIXANTS DES LIGANDS

HYDROPHOBES :

POLYPEPTIDES ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES, ET LEURS APPLICATIONS

<130> D15947

<160> 16

<170> PatentIn Vers. 2.0

<210> 1

<211> 676

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (43)..(552)

<223> cDNA396 (676) /g1 (hOBPIIa-alpha)

<400> 1

```

cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg      54
                                   Met Lys Thr Leu
                                   1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg      102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Leu Ser Phe Thr Leu
  5              10              15              20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc      150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
              25              30              35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg      198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
              40              45              50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc      246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
  55              60              65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg      294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
  70              75              80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac      342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
  85              90              95              100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac      390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
              105              110              115

```



cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggt agg aat 438
 Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Arg Asn
 120 125 130

cct aat acc aac ctg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag 486
 Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
 135 140 145

cac aag gga ctc tcg gag gag gac att ttc atg ccc ctg cag acg gga 534
 His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly
 150 155 160

agc tgc gtt ctc gaa cac taggcagccc ccgggtctgc acctccagag 582
 Ser Cys Val Leu Glu His
 165 170

cccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc 642
 ttccctgctc ccaccacact gactccaaat aaag 676

<210> 2
 <211> 170
 <212> PRT
 <213> Homo sapiens

<400> 2
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15
 Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30
 Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45
 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
 50 55 60
 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80
 Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95
 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110
 Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125
 Val Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys
 130 135 140
 Lys Leu Val Gln His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro
 145 150 155 160
 Leu Gln Thr Gly Ser Cys Val Leu Glu His



165

170

<210> 3
 <211> 725
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (43)..(480)
 <223> cDNA396 (725) /SM12 (hOBPIIa-beta)

<400> 3
 cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
 Met Lys Thr Leu
 1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
 Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
 5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
 Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
 25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
 Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
 40 45 50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc 246
 Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
 55 60 65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg 294
 Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
 70 75 80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac 342
 Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85 90 95 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac 390
 Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
 105 110 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggg ccg tgc 438
 Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Pro Cys
 120 125 130

cgc tgt ccc cac gtc ggc tca cct ggc cac ctc acc tgc agg 480
 Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr Cys Arg
 135 140 145

taggaatcct aataccaacc tggaggccct ggaagaattt aagaaattgg tgcagcacia 540

gggactctcg gaggaggaca ttttcatgcc cctgcagacg ggaagctgcg ttctcgaaca 600



ctaggcagcc cccgggtctg cacctccaga gccacccta ccaccagaca cagagcccgg 660
 accacctgga cctaccctcc agccatgacc ctccctgct cccaccacc tgactccaaa 720
 taaag 725

<210> 4
 <211> 146
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15
 Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30
 Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45
 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
 50 55 60
 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80
 Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95
 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110
 Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125
 Val Gly Pro Cys Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr
 130 135 140
 Cys Arg
 145

<210> 5
 <211> 741
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (43)..(726)
 <223> cDNA396 (741) /SM4 (hOBPIIa-gamma)

<400> 5
 cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
 Met Lys Thr Leu
 1



ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg	102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu	
5 10 15 20	
gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc	150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val	
25 30 35	
gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg	198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val	
40 45 50	
aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc	246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe	
55 60 65	
atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg	294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr	
70 75 80	
gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac	342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr	
85 90 95 100	
ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac	390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp	
105 110 115	
cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg gca tct gct	438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Ala Ser Ala	
120 125 130	
ccc tgc agg gcc gtg ccg ctg tcc cca cgt cgg ctc acc tgg cca cct	486
Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu Thr Trp Pro Pro	
135 140 145	
cac ctg cag gta gga atc cta ata cca acc tgg agg ccc tgg aag aat	534
His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg Pro Trp Lys Asn	
150 155 160	
tta aga aat tgg tgc agc aca agg gac tct cgg agg agg aca ttt tca	582
Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg Arg Thr Phe Ser	
165 170 175 180	
tgc ccc tgc aga cgg gaa gct gcg ttc tcg aac act agg cag ccc ccg	630
Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr Arg Gln Pro Pro	
185 190 195	
ggg ctg cac ctc cag agc cca ccc tac cac cag aca cag agc ccg gac	678
Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr Gln Ser Pro Asp	
200 205 210	
cac ctg gac cta ccc tcc agc cat gac cct tcc ctg ctc cca ccc acc	726
His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu Leu Pro Pro Thr	
215 220 225	
tgactccaaa taaag	741



<210> 6
 <211> 228
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Leu
 1 5 10 15
 Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30
 Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45
 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
 50 55 60
 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80
 Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95
 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110
 Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125
 Val Ala Ser Ala Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu
 130 135 140
 Thr Trp Pro Pro His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg
 145 150 155 160
 Pro Trp Lys Asn Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg
 165 170 175
 Arg Thr Phe Ser Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr
 180 185 190
 Arg Gln Pro Pro Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr
 195 200 205
 Gln Ser Pro Asp His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu
 210 215 220
 Leu Pro Pro Thr
 225

<210> 7
 <211> 607
 <212> ADN
 <213> Homo sapiens



<220>

<221> CDS

<222> (43)..(483)

<223> cDNA396 (607) - forme courte (hOBPIIa-delta)

<400> 7

```

cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg      54
                                     Met Lys Thr Leu
                                     1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg      102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
   5                               10                               15                               20

gag gag gag gat gag gga gaa tcg gtg cat cca gaa gaa aat cct gat      150
Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu Glu Asn Pro Asp
                25                               30                               35

gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa      198
Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
                40                               45                               50

gct cat ata cct gca gga gct gcc cgg gac gga cga cta cgt ctt tta      246
Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg Leu Arg Leu Leu
                55                               60                               65

ctg caa aga cca gcg ccg tgg ggg cct gcg cta cat ggg aaa gct tgt      294
Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His Gly Lys Ala Cys
                70                               75                               80

ggc atc tgc tcc ctg cag ggc cgt gcc gct gtc ccc acc ttg gct cac      342
Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro Thr Leu Ala His
                85                               90                               95

ctg gcc acc tca cct gca ggt agg aat cct aat acc aac ctg gag gcc      390
Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala
                105                               110                               115

ctg gaa gaa ttt aag aaa ttg gtg cag cgc aag gga ctc tcg gag gag      438
Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu
                120                               125                               130

gac att ttc atg ccc ctg cag acg gga agc tgc gtt ctc gaa cac      483
Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val Leu Glu His
                135                               140                               145

taggcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 543

ccacctggac ctaccctcca gccatgaccc ttccctgctc ccaccacct gactccaaat 603

aaag                                                                    607

```

<210> 8

<211> 147

<212> PRT

<213> Homo sapiens



<400> 8

Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu
 20 25 30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
 35 40 45

Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg
 50 55 60

Leu Arg Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His
 65 70 75 80

Gly Lys Ala Cys Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro
 85 90 95

Thr Leu Ala His Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr
 100 105 110

Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly
 115 120 125

Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val
 130 135 140

Leu Glu His
 145

<210> 9

<211> 676

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (43)..(552)

<223> cDNA2098 (676) - forme classique (hOBPIIb-alpha)

<400> 9

cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
 Met Lys Thr Leu
 1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
 Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Leu Ser Phe Thr Leu
 5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
 Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
 25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
 Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val



40	45	50	
aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc acc ttc			246
Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe Thr Phe			
55	60	65	
atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg			294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr			
70	75	80	
gag gag cct ggc aaa tac agc gcc tat ggg ggc agg aag ctc atg tac			342
Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg Lys Leu Met Tyr			
85	90	95	100
ctg cag gag ctg ccc agg agg gac cac tac atc ttt tac tgc aaa gac			390
Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe Tyr Cys Lys Asp			
105	110	115	
cag cac cat ggg ggc ctg ctc cac atg gga aag ctt gtg ggt agg aat			438
Gln His His Gly Gly Leu Leu His Met Gly Lys Leu Val Gly Arg Asn			
120	125	130	
tct gat acc aac cgg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag			486
Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln			
135	140	145	
cgc aag gga ctc tgc gag gag gac att ttc acg ccc ctg cag acg gga			534
Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro Leu Gln Thr Gly			
150	155	160	
agc tgc gtt ccc gaa cac taggcagccc ccgggtctgc acctccagag			582
Ser Cys Val Pro Glu His			
165	170		
ccccccctac caccagacac agagcccgga ccacctggac ctacctcca gccatgaccc			642
ttccctgctc ccacccacct gactccaaat aaag			676
<210> 10			
<211> 170			
<212> PRT			
<213> Homo sapiens			
<400> 10			
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu			
1	5	10	15
Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys			
20	25	30	
Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys			
35	40	45	
Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala			
50	55	60	
Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu			

10

65				70				75				80			
Met	Arg	Lys	Thr	Glu 85	Glu	Pro	Gly	Lys	Tyr 90	Ser	Ala	Tyr	Gly	Gly 95	Arg
Lys	Leu	Met	Tyr 100	Leu	Gln	Glu	Leu	Pro 105	Arg	Arg	Asp	His	Tyr 110	Ile	Phe
Tyr	Cys	Lys 115	Asp	Gln	His	His	Gly 120	Gly	Leu	Leu	His	Met 125	Gly	Lys	Leu
Val	Gly 130	Arg	Asn	Ser	Asp	Thr 135	Asn	Arg	Glu	Ala	Leu 140	Glu	Glu	Phe	Lys
Lys 145	Leu	Val	Gln	Arg	Lys 150	Gly	Leu	Ser	Glu	Glu 155	Asp	Ile	Phe	Thr	Pro 160
Leu	Gln	Thr	Gly	Ser 165	Cys	Val	Pro	Glu	His 170						

```
<210> 11
<211> 782
<212> ADN
<213> Homo sapiens
```

```
<220>
<221> CDS
<222> (1)..(537)
<223> cDNA2098 (782) - forme longue (hOBPIIb-beta)
```

<400>	11															48
cgc	cca	gtg	acc	tgc	cga	ggg	cgg	cag	cac	aga	gct	ctg	gag	atg	aag	
Arg	Pro	Val	Thr	Cys	Arg	Gly	Arg	Gln	His	Arg	Ala	Leu	Glu	Met	Lys	
1				5					10					15		
acc	ctg	ttc	ctg	ggg	gtc	acg	ctc	ggc	ctg	gcc	gct	gcc	ctg	tcc	ttc	96
Thr	Leu	Phe	Leu	Gly	Val	Thr	Leu	Gly	Leu	Ala	Ala	Ala	Leu	Ser	Phe	
			20					25					30			
acc	ctg	gag	gag	gag	gat	atc	aca	ggg	acc	tgg	tac	gtg	aag	gcc	atg	144
Thr	Leu	Glu	Glu	Glu	Asp	Ile	Thr	Gly	Thr	Trp	Tyr	Val	Lys	Ala	Met	
		35					40					45				
gtg	gtc	gat	aag	gac	ttt	ccg	gag	gac	agg	agg	ccc	agg	aag	gtg	tcc	192
Val	Val	Asp	Lys	Asp	Phe	Pro	Glu	Asp	Arg	Arg	Pro	Arg	Lys	Val	Ser	
	50					55					60					
cca	gtg	aag	gtg	aca	gcc	ctg	ggc	ggg	ggg	aag	ttg	gaa	gcc	acg	ttc	240
Pro	Val	Lys	Val	Thr	Ala	Leu	Gly	Gly	Gly	Lys	Leu	Glu	Ala	Thr	Phe	
65					70					75					80	
acc	ttc	atg	agg	gag	gat	cgg	tgc	atc	cag	aag	aaa	atc	ctg	atg	cgg	288
Thr	Phe	Met	Arg	Glu	Asp	Arg	Cys	Ile	Gln	Lys	Lys	Ile	Leu	Met	Arg	
				85					90					95		
aag	acg	gag	gag	cct	ggc	aaa	tac	agc	gcc	tgc	ttg	tcc	gca	gtc	gag	336
Lys	Thr	Glu	Glu	Pro	Gly	Lys	Tyr	Ser	Ala	Cys	Leu	Ser	Ala	Val	Glu	

100	105	110	
atg gac cag atc acg cct gcc ctc tgg gag gcc cta gcc att gac aca			384
Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr			
115	120	125	
ttg agg aag ctg agg att ggg aca agg agg cca agg att aga tgg ggg			432
Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly			
130	135	140	
cag gaa gct cat gta cct gca gga gct gcc cag gag gga cca cta cat			480
Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His			
145	150	155	160
ctt tta ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa			528
Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys			
165	170	175	
gct tgt ggg taggaattct gataccaacc gggaggccct ggaagaattt			577
Ala Cys Gly			
aagaaattgg tgcagcgcaa gggactctcg gaggaggaca ttttcacgcc cctgcagacg			637
ggaagctgcg ttcccgaaca ctaggcagcc cccgggtctg cacctccaga gcccacccta			697
ccaccagaca cagagcccgg accacctgga cctaccctcc agccatgacc cttccctgct			757
cccacccacc tgactccaaa taaag			782
 <210> 12			
<211> 179			
<212> PRT			
<213> Homo sapiens			
 <400> 12			
Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys			
1	5	10	15
Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe			
20	25	30	
Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met			
35	40	45	
Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser			
50	55	60	
Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe			
65	70	75	80
Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg			
85	90	95	
Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu			
100	105	110	
Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr			



cagagcccg accacctgga cctaccctcc agccatgacc cttccctgct cccacccacc 527
 tgactccaaa taaag 542

<210> 14
 <211> 85
 <212> PRT
 <213> Homo sapiens

<400> 14
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Leu
 1 5 10 15
 Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu
 20 25 30
 Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
 35 40 45
 Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro
 50 55 60
 Leu His Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His
 65 70 75 80
 Gly Lys Ala Cys Gly
 85

<210> 15
 <211> 10664
 <212> ADN
 <213> Homo sapiens

<220>
 <223> Gène hOBPIIa

<400> 15
 ctctgcttct atattcggcc tccaaatctg tgctcaatgc agcaactgga gtgaccctt 60
 aaatacgtaa gtcacagctt gcctttgtca gagctatcca gggctcttca ctgagagcag 120
 aagctgaagt cctcgtggtg gtccttaatc cctacatggc cgttccaccc actccccagc 180
 ctcatgtgtg gccggtctcc ctggatcatt tgctgtggct gctctgccgt gtgttcccg 240
 aaactgccag cgttctccca cctctgggct ggcactggat gctcctgcca cctggactcc 300
 tcttccagct gacgagctca tggttgtct ccttcatgtc ttaaattcgg tgtttgaatg 360
 ccacttggc gaggtgttc ctcataat catgtaaagg acaaataacc ccttgtctgg 420
 cactcctgt ccaactgtc cactttgctt tttccatagc actgatcacc atttaaaata 480
 atgtatgaac cagtgaacca tcagtgaacca gagacagtag attgaacagg tattgcttgc 540
 acagtgaaca atcagctcta gaggtttaa gagaaccaca aagaatggct caggattgca 600



cgagggcagt taaggaagaa ataaacaggg gtggggatcc tttccacaag tgtgggttca 660
gaccgcatgg gagaaggtgt ggttctccca agggaagctg gagaagtttg ctgggttgtc 720
ccagccacag ctggcccacg gtcagggcag aggccagcag gcagggaaat gtaggctggg 780
tctggcaagc aggaagcctc tcttcccccg caaggcaggt gggctggggc tgggaagcct 840
acaagagtca ctgggaaccc acaggtgcag atgccactga atctcaatag gaagccatct 900
gggggggtcc cctgaattca atgtggtgtg tgccgccaga cgtccccaac ttgtgccact 960
gccatttata caggaagaga aagaaaaagg aagaaatgga agcatctgga accagtcata 1020
ctggacccat gctaggaggt gcttccccct acgcctcaac caacaaagct tagcattgca 1080
ccaggtgcac aagagaagcg cctgctggct ccagctccat tgcctcgga ccagccatga 1140
agggtgcgtg tggagctgga ggcaagacat tgatagctgg cactgcaatt cacttattta 1200
ttttgttcat ttaagtccc ctgcacctag aatataagcc cccaagcac aggacatttg 1260
ttttattgat cgatgtattc cttgtgcccc aaagaatgag aggcacttag aaagtctgca 1320
aaaatcaaac ataaaaatga acctttattc agtcattggt attttgatgg atatttgagg 1380
catttccaag atttgtaagt aacaattgaa cccttctcct ggttcatgtg tgggagtgt 1440
tctgttcaaa tgatgcttct gagtggagtt gctgagtctt tggtcttagg ttttttttt 1500
ttaagcattt atgctcattg tggtttttaa attaaacatt taacctgag aactgtaga 1560
ttcccatgca attgtaaaaa gccacacaga gctatcatgt gtatcttcac ctggcttgct 1620
ccagcccaa cccagcaat gacagacctg ttctccactc ctgcaatctg ctcatgcaa 1680
gaatgtcgtc tattgcaatc ataaaattgt gggattggct ttttttttcc tgtgcagcat 1740
cattctctgg agattcatcc tattgttgca tttatcaata gtttattcca tttacttct 1800
gagtagtgct ctatggtatg gatgtaccac agtctgttta accattcacc tgttgaggga 1860
agtctgtgtt tatagatttg ggctatgaca catgtatagg ttttgcatg gacatcagtt 1920
ttcattttcc tgggacaaag gccaggggt tctattgctg gattctatgc ttgttacagg 1980
gtcatttttg ttttgttttg ctttgtttta ttttgttttt aacctgtcaa gccattttcc 2040
agatccagtt tccatgcata ctaccaggc ttcagtatga tcactatgat cttatctcag 2100
ccaccttaat aggtatgtac tgatatatca tggcttttat ttgcatttca ctgatgacta 2160
atggtgttga gcatcttttc atgtgtttat ttgccatctg tatatcctct ctagtcaagt 2220
gtcccttcat gtcttttgtt tacgttctat ttttgaaact gttgagtttt gaaaaattct 2280
ttataaagta tagaaactaa ttctttgttg aatatgtagt ttgtcaatat tttctttcag 2340



tcttttggtt gtctttttat tctgttaaca ggggtctctta cagagcaaaa ggttttttatt 2400
ttgatgaagt ctattttaac aatttttctt tttatggatc atatttttgg caacaaatct 2460
aaaacctcct gaccagctc catatgtcaa agattttctt gttttctaaa agttttatag 2520
ttttaagttt tatgttgaag tctatgatcc attttgagtt aattttcata cagggtgtga 2580
gaccaaggtt gtggttcttc tttttctttg gtggttttgt ctgtggatgt ccagttgctc 2640
cagcccattt gctacaaaag ctatctttcc tccactgaat tacttttgca tctttgtaaa 2700
aatgtaattg ggtgtatttg tacaggctctg tttgaggatt ctttattgtc tgtccattg 2760
atctatgcat ctgtccatgt gctagctata taagtcttga aagggtagcc tgtagctggg 2820
tgtggtggcg agagcctgta atcccagcta ttcgggaagt ggaggcagga gaatcgattg 2880
aaccaggag gtggaggttg cagtgagtca agatcgtgcc actgcactcc agcctgggtg 2940
acagagtga ctccatataa aaaaataaaa acaaaaataa agtagcgtga tttctccac 3000
cttattcttt tttaaaaaag ttttagctat tccagttcct ttgcctttcc atataaattt 3060
tagaataatc ttgtctatat ctaccacaat tctttctgga attttgatag aaattttggt 3120
acatctttat attatttgag agaaatgata tttttactat gttgagtctt ctaatccatg 3180
gatataacgt ctctccattt gtttagatct tctttgattt attttataat cattgcattg 3240
ttttcagcat acaaatccag catatgggtt gttagactta tgcctaggca tctcattttt 3300
ttagccatta taaatagtag tgtgttttta agtttagggt ccattactag tatgtagaca 3360
cacaattgat ctttgtatat ttatcttctg tcttgcaacc ttgctgaact cacttactag 3420
ttctaggagg tgtattgttt tcttttggtt tgtttttcaa tttcttgga ttttctacag 3480
agataataat gtcactgca gatgcagttt tcttcttcc tttccaattt gtatgtcttt 3540
aatttccttt ttaaaaaacc tatattactc tgactagaac tttctgtact atgttaaata 3600
caagtggtag agtggacatc cttgccttgt cctgatgtt aaagagaaag catttgtaac 3660
tgagtatccc agcctcaaaa tgtgcttaaa aacttttttc ctttcttgct ttcagccttg 3720
aaacatactt cgaaactctt tatttctccc tttcccacca ggcacttccg tgagcagtgc 3780
tcgcttatct aattatgtgc ttacttagaa attccagggg ccaattttga aacaaaccag 3840
gcagagagac ccagctgcag aatcctccct cttaggggga gttacagga gcctaccact 3900
tcccggctga aatcaggatg acgcaaacca gacctccgga cagacgattg atgactcaca 3960
ataaccatca gaacaagatg cagaccaaca tcctcctgca ccattcccac atatttccca 4020
caccttttcc tccttaaacc ccttcgctca gtccagaaaa tctgaatggt cttttaaagg 4080



catgggtctg gccattcccc aactgccagt atttgaataa agctgctttc cttttaccac 4140
acctcacttc tcatgccttg acttctgagc agcgagcagc tggacttgag ccagttacac 4200
atcagtcttt caccattaga aataatgtag ccataggttt ttttttcgta gatgttcttt 4260
ttcaagttaa agaagttctc ttctattcct atttttctga gaggtgttat cccgaatgag 4320
tggtgaattt tgttaaatac ttttaacaac ccaacaggaa ccaccatcag aagccatcca 4380
gagaaacgca ccaggccaca aagcactggg ggccagggat cttgcccctg ctgtctgcca 4440
tggttgacc cccagcctcc aaccctacca tcccctgacg gtgtctgcag cagttgaacc 4500
caaccagcat ctaaaagaac acagttggtg aacgagactg ggacacaggg caagatgggt 4560
ggacaatggg aggtcctgg agagcaccgc taccagcgag catagaattc atgggggtga 4620
cctgttccct gaagcatctg cgcgtgtgt tccagcattt tcttcaagga ttgagccagc 4680
agcaccagtg tcatacgggtg cttaaataca tgattcacag ccaaccaatg aaatacaagg 4740
tgccggctgg gcgcggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcggg 4800
cggatcacga ggtcaggaga tcgagaccat cccggctaaa acggtgaaac cccgtctcta 4860
ctaaaaatac aaaaaaaatt agccgggcgt agtggcgggc gcctgtagtc ccagctactt 4920
gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg gagcttgagc tgagccaaga 4980
tcccgcact gcactccagc ctgggcgaca gagcgagact ccgtctcaaa aaaaaaaaaa 5040
aaaaaaaaa aaaagaaata caaggtgcct ggggtacagg caacaaaaat ggggaacgga 5100
gaaatcttcc cgttgaggtg gagactccag cctgtttgct catctctggt cctgaagagc 5160
ccagtccccg ccctaaggat ggggtttctg ctcggcagca cttgccgtga gaggggtgag 5220
gcactgggtc acgccagccc tttctttata gccccaggg tttattagag tgtgcattag 5280
tattatttag caagcacttg aggtgtctg atcttgggcc agagaggcac aaagattgtg 5340
tgggcccagc acctgcctgc agagcgtggg tcagccttgg gtcccaggc agatgacgca 5400
ggcccgga ctacagggtc caggaggag aaggcaaagt tctggctcag gttggctggg 5460
gatgaggcca gcggagccag gtgcccaagg gagctcagcc acaaactctg agcacaggct 5520
ggcaggtggc tcttgatgct catgccccc atttatctaa agggatgaga ttcaaggcct 5580
gtcctggtgg ctgggggccc tcgaagctga cgagagaggg gatgtagagt gaatgtatat 5640
tccactctac cacttgatt taacaggag atggatgatg aacacgtgca ggaggaaaca 5700
ggcaggacaa tccagagaga tcacgtgttc tgaggacagc acagccaggc tccggtacgg 5760
agtgaagcgg ggtggggcag gcggcggggt ccctcatatg gcccgaggag gccgtatata 5820

tactgacctt gagccacaca atagtgcctt tctctgcccc taggaagctc gagtgcaggg 5880
tccaggtggg gaaaatcaat gcagagtggg tcccagagtg ggcggaagct tgggctctag 5940
ggcgtgcggg actcagctgg cagcagcccc acatctctat ggttctaaag ccaggtccat 6000
ttctgctcag cagggagatg gccagttccc cagaggactt gcccaggggc ccagccgtgg 6060
ccctgggagg gctcaggggt gagggaggga gaattccgag ccgtggggccc tccctctgtg 6120
gctcgggggtg cagacgtgca ctaccctgcc ctgtcctctg gagtcctccg ggctcatcca 6180
agtggggcga tttgggtgca aaccctggac agcatcgtgt gtcttttctt ctctggctgg 6240
cactgaattg ctgttcacca gatgccagca atgaggacac tccccctccc ggatgcagga 6300
ctggttgcca ccatggggag ggggtgcaccg tgccacgtgc tcccaggaca ctggccaggg 6360
ggctataaag aacatctcga gaggagccag cacagccttg ttcagacgcc cagtgcactg 6420
ccgaggtcgg cagcacagag ctctggagat gaagaccctg ttcctgggtg tcacgctcgg 6480
cctggccgct gccctgtcct tcaccctgga ggaggaggat gtgagctggg ttggcgtggg 6540
cggatggagg agccaggtgg actcctgggc agggggcagt gccagggggc ctgctttagg 6600
aggtgtcact taagcctggg gtgggtgggtg aggggtccta ttgttccttc agcagacaat 6660
gctcccatg agggccaggg ccggagcagg ctcggtcag gggttcctgc tgcactgacg 6720
cctgaagccc gaaggtctcg cagggttggg ccctgtggag ggagggtca cctggtgctg 6780
gggcccggg gtccatggg tgacagacatg ccctccttcc actgggggct gggagccctg 6840
agcagggggc tggctctaac tcaactccagc tgagctctaa ctaaggtgca ggaaccagc 6900
ctgcctttag ggggtggcagc cgggcacccat ggggtgtctg ttatagctgc aggcctgagt 6960
gccaggggtca gagtagaate tgggccaccc atggtgggct cacggccttg gcctgctcca 7020
gatcacaggg acctggtacg tgaaggccat ggtggtcgat aaggactttc cggaggacag 7080
gaggcccagg aaggtgtccc cagtgaaggt gacagccctg ggcggtggga acttgaagc 7140
cacgttcacc ttcattgtgag tgttgccac tgcagggccc ctcaggccac ttctgctccc 7200
cgccccagac ccacctggtg cccattgccc catccacatt tcgggtgttg ggaagagtca 7260
ccccctgcct tggagggaaa cagccagggc atcctgaagc tcggtggggg gggggggcag 7320
tggaattttc aggttgccgg gtcagggcca tgcaccaggt gagctgagga tgggccaggt 7380
gtgtcctggg agccgctgcc cgcgtgtctc ctgttttcca ggagggagga tcggtgcatc 7440
cagaagaaaa tcctgatgcy gaagacggag gagcctggca aatacagcgc ctgtgagccc 7500
ctccccgacc cccactccc catgcccac cccggatgca ccagccccac tgcaggtgga 7560

gagtgccag gccacacttc tgccaggggc ccagccctgc ccacctcaa ggaggggctg 7620
gcctctcctt cctggggggc tggtagccct gacatcagac accaggtgtg acaggcttgt 7680
ccacagtaga gatggaccag atcaagcctg ccctctggga ggcctagcc attgacacat 7740
tgaggaatcc gagtgttagg gaccaggagg ccgaggggta gggatgggaa gccaaaggctg 7800
agggtttggg atcaggaatc cgagggttag ggacagggaa gtggggcagg agcagctgct 7860
ggagctggga aggccggact ctagtcctgg acgtgctctg gccttggtggc tccattactt 7920
gcattgggac cttccgagag gaggtcctg cctccgtgtc cgggtccatg ctgtgcggag 7980
cagccaggcc tggctcaggc tgtccagggc acctgggtga cactgaaac attcctgagt 8040
gtttcttcgt gtggtcctga gtgctctctc cggaatgag ggcactgaag acctatcttc 8100
tctgtcatct acagatgggg gcaggaagct catatactg caggagctgc ccgggacgga 8160
cgactacgtc ttttactgca aagaccagcg ccgtgggggc ctgcgctaca tgggaaagct 8220
tgtgggtgag gggcccgtg ggcctgcat gtctgcccc atggtctctg cctccagaag 8280
ccagtgaac caccatcatc acgccctggc acggggggaa aaggaagccc cctgcgccgg 8340
ccttcgtgtg ctaggcacca agcgtgccc tggatggctg gtccaagttc ctgaagtggg 8400
agtggggtgg gccaggcagg gacagacacg gccctcgggtg acgtgaacct gccaagggcc 8460
gcttggtggg tctcaggtgt aggggcctca ccttaagggg gaggtancat cttacagag 8520
ctcttcattg ggcagggact ctccaggcg gcagggcagc cagtgcctct gggacacaag 8580
gtccctccag gtgaggggtg tgacctgca gagtggcttt gggagctgcc caggtcccc 8640
tgggggtgct gagtggcttg gacctgcca ctgtcccctt tcctggggac ctctcacctg 8700
ggcgggtggc gtctcctctg tccccagtc caccctgag ctcttgcca ttctcaggcc 8760
tcctctcccc cttgcctggt gctggacagt tgccatctct tctgtcccca gccccacccc 8820
tgagctctga tccactctcg ggctctccc ccgtcctgat gctgggccgt ggtcgtctcc 8880
ttttagcatc tgctccctgc agggccgtgc cgctgtcccc acgtcggctc acctggccac 8940
ctcacctgca ggtaggaatc ctaataccaa cctggaggcc ctggaagaat ttaagaaatt 9000
ggtgcagcac aagggactct cggaggagga cattttcacg cccctgcaga cgggtgagga 9060
cggctgtgcc cagtaccccg tgttcccctg tgtctctgtg tgatctccag tgtcccatga 9120
ccctcgtgtc ctcccatgtc ccccgcatc cccatgtgcc ccgagtctcc tcgcaggggc 9180
tctgggccct gcttagcatc ctctcgttg gagggctctg actctgggct gcgatggggt 9240
ctggggctcc gcgctctggg ctgcgatgcg gtctggggct ccgcgctctg ggctgcgatg 9300

gggctctgggg ctccgcgctc tgggctgcga tggggctctgg ggctccgcgc tctgggctgc 9360
 gatgggggtct ggggctccgc gctctgggct gcgatgggggt ctggggctcc gcgctctggg 9420
 ctgcgatggg gtctggggct ccgcgctctg ggctgcgatg gggctctgggg ctccgcgctc 9480
 tgggctgcga tggggctctgg ggctccgcgc tctgggctgc gatgggggtct ggggctctga 9540
 gctctggggt gcgatgggggt ctgggccctg gtctagggcg ccttctaata cctggggtttt 9600
 tcttgggtctc tgcaggaagc tgcgttctcg aacactaggg tgagtgagcc tttaggaggg 9660
 cactggacaa gccagagtc ctgggttccc ggggttcgag ggtacatctg ctctggccct 9720
 tcccatccca cacagccagg gagaccccc cagggtcagg cacgagggtg gcacctcaga 9780
 gtctgcccac ccaaaattcc tgggacattc gggaagtctt ttgttttacc attcctgcac 9840
 ctgccagccc gagtgagggt cctcctcggc ctttccacag cgaggcctcc ctccggctcc 9900
 ctcagggtgtc agctcgcccc atcgccccct gcacctgtcc cactcggtc tactcccccc 9960
 caccactca ccaaggatgc tcagaggcct gccagtcac tggaggagct gaggctgtc 10020
 tggagccccg aggctgccc gcagtggccg atgtggaagc tgagagcct ggggagggag 10080
 ctctgggcct ggctctgcc ctaccctgca gcctccctga cctctgctcc tcttcccagc 10140
 acagcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccga 10200
 ccacctggac ctaccctcca gccatgacct ttccctgctc ccaccacct gactccaaat 10260
 aaagagcttc tccccagct ctgggcaggc ctatctgtgg ggacggaggg gctcgcaccg 10320
 gctccctggg aggcctgctg ggaggggggag ccacaaggga agctggggag atgttgacc 10380
 tagcaagggc caaaccacaa gaggcacgat gtccaacagg ctgtgggggt gcgtgatgtc 10440
 ctggaggggc ctccggaaga gccgccctcg gatctcaggc tcagcttggg acgggcaggg 10500
 ctctgggcag gagcacttgg gaagccactg ggagggccgg agtggggaca cggcgtcaga 10560
 cctgcatcag tgggtccacc acagggtcc cgccccggg gtctcagtgg gatcctccga 10620
 gcgggtccac tttancatcc tnggtcagt ggcactcgt gata 10664

<210> 16

<211> 13591

<212> ADN

<213> Homo sapiens

<220>

<223> Gène hOBPIIb

<400> 16

tctctgcttc tatattcggc ctccaaatct gtgctcaatg cagcaactgg agtgaccct 60



taaatacgta agtcacagct tgcctttgtc agagctctcc agggctcttc actcagagca 120
gaagctgaag tctcgtggt ggtccttaat ccctacatgg cagttccacc cactccccag 180
cctcatgtgt ggccggtctc cctggatcat ttgctgtggc tgctctgctg tgtgttcccg 240
gaaactgcc a gcattctccc acctccaggc tggcactgga tgctcctgcc acctggactc 300
ctcttccagc tgacgagctc atggcttgct tccttcatgt cttaaattcg gtgtttgaat 360
gccaccttgg cggggctggt cctcatcaat tcatttaaag gacaaataac cccttgtcct 420
gacactcctg tccaacttgt ccactttgct ttttccatag cactgatcac catttaaaat 480
aacgtatgaa ccagtgaccc atcaggatcc agagacaata gattgaacag ggattgcttg 540
cacaatgaac aatcagctct agagggttaa agagaaccac aaagaatggc tcaggattgc 600
acgagggcag ttaaggaaga aataaacagg ggtggggatc ctttccacaa gtctgggttc 660
agacctcatg ggagaagggt tggttctccc aagggaagct ggagaagttt gctgggttgt 720
cccagccaca gctggccccc ggtcagggca gaggccagca ggcagggaaa tgtaggctgg 780
gtctggcaag caggaagcct ctctccccac ccaaggcagg tgggctgggg ctgggaagcc 840
tgcaagagtc actgagaacc cacagggtgca gatgccactg aatctcaata ggaagccatc 900
tgggggcggt cccctgaatt caatgtggtg tgtgccgcca gacgtccccg acttgtgcca 960
ctgccatttg cataggaaga gaaagaaaaa ggaagaaatg gaagcatctg gaaccagtca 1020
tcttggaacc atgctaggag gcgcttctcc ctacgcctca accaacaag cttagcattg 1080
caccagggtgc acaagagaag cgcttgcttg ctccagctcc attgcctcgg aaccggccat 1140
gaagggtacg tgtggagctg gaggcaagac attgatagct ggcactgcaa ttcacttatt 1200
tattgtgttc attttaagtc ccctgcacct agaataaag ccccccgagc acaggacatt 1260
tgtttcattg atcgatgtat tccttgtgcc ccaaagaatg agaggcatct agaaagtctg 1320
caaaaatcaa acataaaaat gaacctttat tcagtcattg ttattttgat ggaaatttga 1380
ggcatttcca agatttgtaa gtaacaattg aacccttctg ctggttcatg tgtgggagtg 1440
tatctgttga aatgatgctt ctgagtggag ttgctgagtc tttggctcta ggtttttttt 1500
ttttaagcat ttatgcttat tgtggttttt aaattaaaca tttaaccctg agacattgta 1560
gattccccat cagttgtaaa aagccacaca gagctatcat gtgtatcttc acctggcttg 1620
ctccagcccc aaccccagca atgacagacc tgttctccac tcctgaaatc tgctcattgc 1680
aagaatgtcg tctattgcaa tcataaaatt gtgggattgg cttttttttt tcctgtgcag 1740
catcattctc tggagattca tcccattggt gcatttatca atagtttatt ccattttact 1800



tctgagtagt gctctatggt atggatgtac cacagtctgt ttaactattc acctgttgga 1860
ggatgtctgt gtttatagat ttgggctatg acacatgtac aggtttttgc atggacatca 1920
gttttcattt ttctgggaca aaggcccagg ggttctattg cggggttcta tgcttggtgc 1980
agggtttttt tttttaaac tgtcaagcca tttccagaa actgtccagt ttccatgcat 2040
cctcaccagg cttcagtatg atcactatga tcttatctca gccaccttaa taggtatgta 2100
ctgatataatc atggctttta tttgcatttc actgatgact aatgggtgtg agcatctttt 2160
catgtgttta tttgccatct gtatatctc tctagtcaag tgtcccttca tgtcttttgt 2220
ttacattcta tttttgaaac tgttgagttt tgaaaaattc tttataaagt atagaaaacta 2280
attctttgtt gaatatgtag tttgtcaata tttcttttca gtctttggct tgtcttttta 2340
ttctgttaac aggggtctct acagagcaaa aggtttttat tttgatgaag tctattttta 2400
caatttttcc ttttatggat catatttttg gtgacacatc taaaatctcc tgaccagct 2460
ccatatccca aagattttct ttctgttttc taaagctttt atagttttaa gttttatggt 2520
taagtctatg atccattttg agttaatttt cttgcagggt gtgagaccaa gggtgtggtt 2580
cttctttttg tttggtggtt ttgtctgtgg atgtccagtt aatccagccc atttgctaca 2640
aaagctatct tccattgaat tgcttttgca tctttgtaaa aacttaactg ggtatatgctg 2700
tgcaggctctg tttgaggatt ctttattgtc tgtccattg atctatgcat ctgtccatct 2760
gctagctata taatgagtct tgaaagggta gcctgtagct gggtatggtg gtgtgcatct 2820
gtaatcccag ctacttgga ggcgaggca ggagaatcac ttgaatctgg gaggtggagg 2880
ttgcaatgag tcgagattgt gctactgcac tccagcctgg gtgacagagc aagactctgt 2940
ctctaaaaaa aagaaagaag aagaaaagga aagaaagggg agcctgattc ctcccacttt 3000
attcttaaaa aaaaaagttt tagttattct agttcctttg cctttccata taaatgttag 3060
aataatcttg tctatgtcta caaaaattcc ttctggaatt ttgatagaaa ttgtgttaaa 3120
tctttatatt atttgagaga aatgacgctt ttattatggt gagtctccta atccatggat 3180
atagcacatc tctccatttg tttagatctt ctttgattta ttttataatc attgcattat 3240
tttcagcata caaatccagc atatggggtg ttaaaacttat gcctaggtat ctcttttttt 3300
tagccactat aagtagtatt gtgtttttta gtttaggggtc cgttactagt atgtagacac 3360
acaattgatc tttgtatatt tatcttgat cctgcaacct tgctgaactc gcttaatagt 3420
tctaggaggt gtattgtttt cttttgtttt gtttttcagt ttcttgggat tttctacaga 3480
gacaatcatg tcatctgcag atgcagacag ttttcttctt tcctttccaa tttgtatgtc 3540



tttaatttcc tttttaaaaa acctttattg ctctgactag aactttctgt actatgttaa 3600
atacaagtgg tgaaagtgga catccttgcc ttgtccctga tgttaaggag aaagcgtttg 3660
taactgagta tcccagcctc aaaatgtgct taaaaacttt tttcctttct tgctttcagc 3720
cttgaaacat acttcgaaac tctttatttc tccctttccc accaggcact tctgtgagca 3780
gtgctcgctt atctaattat gtgcttactt agaaattcca ggggccaatt ttgaaacaaa 3840
ccaggcagag agaccagct gcagaatcct ccctcttagg gggagttaca ggtagcctac 3900
cacttcccggt ctgaaatcag gatgactcaa accagacctc tggacagacg attaatgact 3960
catgataacc attggaacaa gatgcagacc aacatcctcc tgcaccattc ccacatattt 4020
cccacacctt ttcctcctta aacccttcg ctcagtcag aaagtttgaa tggcttttta 4080
aaggcatggg gcctggccat tctctactg ctagcatttg aataatgctg ctttcccttc 4140
accacacttc acttctcatg ccttgacttc tgagcagcga gcagctggac ttgagccagt 4200
tacacatcag tctttcacca tgagaaatga tgttagccat aggttttttg tagatgctct 4260
ttgtcaagtt aaggaggttc tcttctattc ctacttttct gagagtgtgt ttcctgaatg 4320
cgtgatgaat ttagtcaagt actttttctg cattgatcga tatgatcatg taatttttct 4380
tcttaacata ttaatatggt tgatttttga gtattgtacc aggtttccat ccctggaata 4440
aactgctttt ggatcatggtg tagaattaat tttatatatt attatttgct aatattttgt 4500
aaaggatttc tgtatctata ttcagaggt atattgttct gcagctttcc tttttgtgtt 4560
gtcagggtttt ggttgggaga tattctctca tcttctgaga ttatgcagag ttggtgttat 4620
ttatttttta aatggttgggt agaattttct aatgaaacca tatggacatg aagaattatt 4680
tttggagct tttaaaatta aaatgttaac ttttaattgtc atagagctat taaagttatc 4740
tatcttatat tggcattgtg ttttttgaga aattggttca tttcacctat gttgtcaa 4800
atatgtggga aggtctcttc gtagtattcc cttattatcc tttgatgtct gcagggtctg 4860
tagtgatagt ctctgttttg ttccagatgt tggttatttg tgtctttctt tctttttttg 4920
tttgtcagtc tttctagaac tgcctcaatt ttattaattt ttccagagaa ctaacccttg 4980
tttcattgtt tttctctgtt gtttttctgc tttcgatttt ttctgcttct tgcttttgat 5040
ttattaattg ctgctcttat ctttgttatt tcttttcttc ttgctttggg tttattttgc 5100
tcttcttttt ctaggttctt gaagtgggag cttagatagt tgattttgga cttctctttt 5160
ctaatatatg catttaatgc ctctcagcac tgatttagct gtgtctcaca aattttgata 5220
tgtttttggt tgtttgggtt tgaggtagag tctcgctctg tatcccaggc tggagtgcag 5280



tggcactatc ttggetcact gcaacctcca cctcctgggt tgaagtaatt ctcctgcttc 5340
agccccctaa gtagctggga ttacaggtat acgccaccat gccagcaat ttattttattt 5400
attttttttt gtatttttag tagagatgag gttttgtcat gctgtccagg ctgggtctcga 5460
actcctgacc tcaagtgate tgcccacgtg gggctccaaa gactgagat tataggcata 5520
agccactgtg ccagcctgt ttttgttttt tttgagatgg agtcttgctc tgtcaccag 5580
gctggagtgc agtggcatga tagccactg caacctccgc ctcccgggtt caagtgattc 5640
tcctgctca gctcccgat tagctgggat tacaggcgtg caccaccaca ctcagctagt 5700
ttttgcattt ttggtagaga tgggttttcg ccatgttggc caggctggtc tcgaattcct 5760
gatcgcaagt gatccaccg cctgggcctc ccaaagtgt ggaattatag gcacaagcca 5820
tcacgcccag ctgatatgtt gtattttcat tttcatccag cctagtatat ttttaaattt 5880
accttcaggc ttcctctttg acctgtggat tattcacagg tgtgttgttt ccaagcatct 5940
gaagaaagat tttccattat ctttctggca ttgatttata gttagattcc atgtggtcag 6000
agactaccct ctatatcatt taaattttta aaaatttctt gaggtcatt ttctcatcca 6060
ggatatggtc tatcttggtat tatattctgg tgcacttgaa aagaatgtgt gtagtgctgt 6120
tgccaggtgg tgtgttctac aaatgtcaat tcgattatgt taattgatgg gtggccgagt 6180
tctttgatgt cctgctgggt tttctgtcta gttgctctga gagaagggtat ttgaagtctc 6240
caactataat tgtggctttt taaatttctc ctttcagttc tgttttgctt cacatattgt 6300
gcagctaattg tttgggtgcac acacgttttag gattactatg tcttctttgc agaattggccc 6360
acttattatt gtataatttc cctctctgtt aattagtctt tgttctaaag tctattttac 6420
ctgttattga aataatgctt ctgctttctt ttgattaatg tttatatatc ttttttcatg 6480
tttttacttt caacctgtct ctattgttac atttgaagta agtttcttgt agacggcata 6540
tagttgtata atggctttta attcactttg ccaatcactg tcatttaata tttaaacat 6600
ttacatttaa tataattatt aacatgctag agtttgtatt atttttatag ggttgctata 6660
acaaaatact accaactggg tggcttagaa caaaaattta ttttctcaca ttctgggtggc 6720
cagaagtctg agatcaagtg agatttgga ggtagaaaaa ccaacaaaat atgttaatga 6780
tgttaccatt gtgggcaact gggactcagt ccttctgagg gccatctgag aaatagggtg 6840
gaaccatatt tgattgttcc acttagaagg aagagtccag accatttatc ccatcagtta 6900
ctatccctac agactgagga tgctgattca cttgcacatc tgggtttcat cagtgtcagg 6960
ggaagcacag aagaaaaggt gcaggcactt aagggtggga gctgtgaata cgtccgtgca 7020



ccaggaacga tctacaactg tggcaggtgg gccgacagcg tgggtcacca gtgccttcag 7080
ccacccttg cactgcactg acccactcac acagcaaadc aagtccattt tgtaactgat 7140
gcttcagcat ggtggccagc caccgtctct gaaataatta aaataggtgg ttggtggcac 7200
aggcgatggt ccctactgtg cagttggtct tgcagcctta actgagattc atcgtaagct 7260
cccactctgc tccttcttgg cgtgcctcag tcccagccag cacatgctct ggtctaggtg 7320
cttgtttggt ggtgccatct aaaccttctt atgcccaggt gtcgagtaaa gccagacac 7380
tgtccacttc agtgaagcac ctaggaatta gcagccctga aaagctgtga tactgtgag 7440
attcaaagaa aagggtcgtg aggaagtgg tgagagagaa agttgtgtgg tagcggcagc 7500
ccaacaggag ccaccatcag aaggaacaca gccatccaga gaagtgtccc aggccacaaa 7560
gcactggggg ccagagatct caccctcat gtctgccatg ggttgacccc caacctccag 7620
ccctaccatc tcctgacggt gtctgcacca gttgaacca accagtatct aaaaggacac 7680
agttggcgaa tgagactggg acacagggca agatgggtgg atgatgggag actcctggag 7740
agcaccgctc gcactgagcg tggacctcat gggagagccc tgttctctga agcatctgtg 7800
catgtcgttc cagcatcttc ttcaaggatt gagccagcag caccagtgtt gtaaggtgct 7860
tagatcaatg attcacaac aaccaatgaa atacgaggtg cctggggtac agtcaacaaa 7920
cacagagaac agagaaacct tcccgttgat gtagagactc cagcctgttt gctcatctct 7980
ggtcctgaag agcccaggcc ccattcttct ctgcccgggg ggatgagggt ggggggttct 8040
tgcttggcag cacttgccgt gggaggggtg agggactggg tcacgccagc ccgttcgtta 8100
cagccccagg gtttattaga gtgtgcatta gtattattga gtaagtactt gagagtgtct 8160
gatcttgggc cagagaggca caaagatgcc attgtgtggg tccagcacct gcctgcagag 8220
cgtgggtcag ccttgggtcc cagggaagat gacacacacc tgggaaatgc agggtcggg 8280
agggagaagg caaagttctg gctcaggttg gctggggatg aggccagcag agccaggtgc 8340
ccaagggacc tcagccacga actctgagca caggctggca ggtgactctt ggtgcctcat 8400
gcgacttggt tatctaaagg gatgagattc aaggcctgtc ctgggggctg ggggccgtca 8460
aagctgacga gagaggggaa tgtatattct aacagggaga tggtcgatga atacgtgcag 8520
gaagaaatgg acaggacaat ccagagagat cgcgtgttct gaggacagca cagccaggct 8580
ccggtacgga gtgaagcggg gtcggggagg cggcggggtc cctcatatgg cccgaggagg 8640
ccgtatataa actgaccctg agccacaaaa tagtgcctc ctctgccct aggaagctcg 8700
agtgcagggt ccagggtggg aaaatcaatg cagagtgggt ccagagtgg gcggaagctt 8760



gggctctagg gcgtgcggga ctcagctggc agcagcccca cgtctctctg gttctaaagc 8820
ccagtccatt tctgctcagc agggagatgg ccagttcccc agaggacttg cccaggggtcc 8880
cagccgtggc cctgggaggg tgcaggggtt ggtgaggag tatcccgagg ctgtgggccc 8940
tgcctctgtg gcttgggggtg cagatggaca ccacctgccc ctgtcctctc gagtcctctg 9000
ggctcatctg agtgggocgcg ttcgggtggt gcaaactctg gatggcatcg tgtgtctttt 9060
cttctctggc tggcactgac ttgctgttca ccagatgcca gcaatgagga cactccccct 9120
cccggatgca ggactgggtg ccaccatggg gaggggtgca ccgtgccacg tgctcccagg 9180
acactggcca gggggctata aagaacatct cgagaggagc cagcacagcc ttgttcagac 9240
gcccagtgac ctgccgaggt cggcagcaca gagctctgga gatgaagacc ctgttcctgg 9300
gtgtcacgct cggcctggcc gctgccctgt ccttcacctt ggaggaggag gatgtgagct 9360
gggttggcgt gggcggatgg aggagccagg cggactcctg ggcagggggc agtgccaggg 9420
gccctgcttt aggaggtgtc actgaaggct ggcttctgac cccgtgctcc cagcctgggg 9480
tggtggtgga ggggtcctat tgttcctcca gcagacacgg cccccgagg cccagggccg 9540
gagcaggctc gtctcagggg ttcctgctgc actgacgcct gaagcccgaa ggtctcgag 9600
ggttgggccc tgtggaggga gggctcacct ggtgctgggg cccgggggtc catgggggtg 9660
agacatgccc tccttccact gggggctggg agccctgagc agggggctgg ctctaactca 9720
ctccagctga gctctaacta aggtgcagga acacagcctg cctttaaggg cagcagccgg 9780
gcaccatggg tgtctgggta tagctgcagg cctgagtgcc agggtcagag tagaatctgg 9840
gccacccatg gtgggctcac ggccttggcc tgctccagat cacagggacc tggtagctga 9900
aggccatggt ggtcgataag gactttccgg aggacaggag gccaggaag gtgtccccag 9960
tgaaggtgac agccctgggc ggtgggaagt tggaagccac gttcaccttc atgtgagtgt 10020
tgcccactgc agggccctc aggccacttt cgctcccctc cccagacca cctggtgccc 10080
attgccccat ccacgtttcg ggtgttgga agagtcaccc cctgccttgg agggaaacag 10140
cctgggcatc ctgaagctcg gtgggggtgg ggacagtgga attttcaggt tgccgggtca 10200
gggccatgca ccaggtgagc tgaggatggg ccatgggtgt cctgggagcc gctgcccgcg 10260
tgtctcctgt tttccaggag ggaggatcgg tgcattcaga agaaaatcct gatgcggaag 10320
acggaggagc ctggcaaata cagcgctgt gagccctcc cccactccca cccccacct 10380
ccccaccgc caacccagc gcaccagcct ccacaggtag agagtgccca ggctgcctt 10440
ttgccagggc cccagctctg cccacctcca aggaggggct ggctctcct tcctgggggg 10500



ctggtggccc tgacatcaga caccgggtgt gacaggcttg tccgcagtcg agatggacca 10560
gatcacgcct gccctctggg aggccctagc cattgacaca ttgaggaagc tgaggattgg 10620
gacaaggagg ccaaggatta ggtacgggga ggctgagggg tagggatggg gaagctgagg 10680
gttaaggatg ggggaagctga ggggtaagga tggggaggct gaggggttagg gacgggggcc 10740
aagggttagg gatcgggaag ctgaggggta gggatgagga ggccgagggt tagggatggg 10800
gaggccgagg gttagtgtct cggaagctga ggattagaga tggggaggct gagggggaag 10860
atgaggggta gggacaggga agtggggcag gagcagctgc tggagctggg aaggccggac 10920
tctagtctct gacgtgtctt ggccttgtag ctccattact tgcattggga ccttcctgag 10980
aggaggctcc tgcctccgtg tccgggtcca cactgtgcgg agcagccagg cctggctcag 11040
gctgtccagg gcacctgggt gaccgctgaa acattcccga gtgtttcttc atgtggcccc 11100
gagtgttctc tctgggaatg acccattttc tctgtcatct acagatgggg gcaggaagct 11160
catgtacctg caggagctgc ccaggaggga ccactacatc ttttactgca aagaccagca 11220
ccatgggggc ctgctccaca tgggaaagct tgtgggtgag gggcttgctg gggctctgct 11280
gtcctgcccc acggtctctg cctctggaag ccggtggaac cagcaccacc atgccctggc 11340
atgggggaaa aggaagcccc ctgtgccggc tttcatgtgc cgggcaccgc tgggcagccg 11400
gtccaagtac ctgaagtggg aatgggatgg gccaggcagg gacagacatg gccctcagt 11460
acatgaacct gccaaaggcc acttctgggg tctcgggtgt aggagggtca ccttaagggg 11520
gagggtcacc ttaagcgctc taagagagct cttcatgggg cagggaccct ccagggcagg 11580
agggtggccg gtgcctctgg gagacaaggt ccctccagggt gagggctgtg accctgcagg 11640
gtggcattgg gagctgcccc ggtcccctgg ggttgccgag tggttgagg agtccccgcc 11700
actgtcccc ttctggggga cctctcatct gggcagtgac tgtctcctct gtccccagtc 11760
ccaccctga gctctcatcc attctcaagt ctctctctct gcttgctcgg tgetgggcta 11820
tggccatctc ttctgtcccc ggccccacc cagagctctg gtccactctt gggctctctc 11880
ccttgctctg gtgtgggca gtggccatct cctctgtccc cagccccacc cccgagctct 11940
gatccactct caggcctctc cccctgtcct ggtgctgggc cgtggctcgtc tcctttcggc 12000
atctgtctcc tgcagggccg tgccgctgtc cccacgtcgg ctccactggc cacctcacct 12060
gcaggtagga attctgatac caaccgggag gccctggaag aatttaagaa attggtgcag 12120
cgcaagggac tctcggagga ggacattttc acgcccctgc agacgggtga ggatggctgt 12180
gcccgatccc ctgtgtccct ctgctgtgtc tgtctgtat ctccagtgtc ccatgacccc 12240



catgtcctcc catgtccccc gcattecccca tgtgccccga gtctcctcgc aggggctccc 12300
gggccctggt tagcgtcctc ctcattggag gctctgtgct ctgggctgcg atgggggtctg 12360
gggctccgcg ctctgggctg cgatggggtc tggggctccg cactctgggc tgcgatgggg 12420
tctggggctc cgcgctctgg gctgcgatgg gctctggggc tctgagctct gggctgtgat 12480
ggggctctggg ccctgggtcta gggcgccttc taatccctgg gtttttcttg gtctctgcag 12540
gaagctgcgt tcccgaacac tagggtgagt gagcctttag gagggcactg gacaagccca 12600
gagtcctggg ttcccggggg tcgagggtac atctgctctg gccctccca tcccacacag 12660
ccaggagagac ccccccaggg tcaggcacga gggtggcacc tcagagtctg cccaccaca 12720
gttcctggga cattcgggaa gtcccttggt ttatcattcc tgcacctgcc agcccgagtg 12780
agggtcctcc tcggccttcc cacagcgagg cctccctccg gctccctcag gtgtcagctc 12840
agcccatcgc cccctgcacc tgtccccacg cggtcactc cccaccatcc accaccaag 12900
gatgctcaga ggctgccca gtcattggag gagctgaggt cgctctggag ccccgagggt 12960
gcccagcagt ggccgatgtg ggggctgcag agcctgggga gggagctctg gccctggcct 13020
ctgccctacc ctgcagcctc cctgacctct gctcctcttc ccagcacagc agcccccg 13080
tctgcacctc cagagccac cctaccacca gacacagagc ccggaccacc tggacctacc 13140
ctccagccat gacccttccc tgctcccacc cacctgactc caaataaagt ctttctccc 13200
cagctctggg caggcctatc tgtggggaca gaggggcttg caccgctcc ctgggaggtc 13260
tgctgggagg gggagccaca agggaagctg gggagatgtt ggacctagca agggccaaac 13320
cacaagaggc acgatgtcca acaggcctga ggggctgcgt gatgtcctgg aggggcctcc 13380
ggaagagccg ccctcaggtc tcaggttcag cttaggacag gcagggccct gggcaggagc 13440
gtttggaaag ccactgggga ggacaggagc ggggacacgg cgtcagggct gcagcagtgg 13500
ggccaccgca gggctcccgc ctcaggggtc tcagagggat cctccaagct ccccatcttt 13560
agcagcctgg gctcagtggc agcactggag a 13591

